

FUNDAÇÃO UNIVERSIDADE FEDERAL DE RONDÔNIA
CAMPUS JI-PARANÁ
DEPARTAMENTO DE MATEMÁTICA E ESTATÍSTICA

PÁBLO DIAS VIEIRA

**COMPARAÇÃO ENTRE CADEIAS DE MARKOV DE PRIMEIRA E SEGUNDA
ORDEM PARA OCORRÊNCIA DE PERÍODOS SECOS E CHUVOSOS NA
REGIÃO AMAZÔNICA.**

JI-PARANÁ, RO
2017

PÁBLO DIAS VIEIRA

**COMPARAÇÃO ENTRE CADEIAS DE MARKOV DE PRIMEIRA E SEGUNDA
ORDEM PARA OCORRÊNCIA DE PERÍODOS SECOS E CHUVOSOS NA
REGIÃO AMAZÔNICA.**

Trabalho de Conclusão de Curso apresentado ao Departamento de Matemática e Estatística, da Fundação Universidade Federal de Rondônia (UNIR) *Campus* de Ji-Paraná como parte dos requisitos para obtenção do título de Bacharel em Estatística.

Orientador: Dr. Dilson Henrique Ramos Evangelista

**JI-PARANÁ, RO
2017**

Dados Internacionais de Catalogação na Publicação
Fundação Universidade Federal de Rondônia
Gerada automaticamente mediante informações fornecidas pelo(a) autor(a)

D541c Dias Vieira, Páblo.

Comparação entre Cadeias de Markov de Primeira e Segunda Ordem para Ocorrência de Períodos Secos e Chuvosos na Região Amazônica / Páblo Dias Vieira. -- Ji-Paraná, RO, 2017.

48 f. : il.

Orientador(a): Prof. Dr. Dilson Henrique Ramos Evangelista

Trabalho de Conclusão de Curso (Graduação em Estatística) - Fundação Universidade Federal de Rondônia

1. Cadeia de Markov. 2. Distribuições de Probabilidades. 3. Ocorrência de Chuva. I. Evangelista, Dilson Henrique Ramos. II. Título.

CDU 519.2



ATA DA SESSÃO DE APRESENTAÇÃO E DEFESA DO TRABALHO DE CONCLUSÃO DE CURSO

Aos 25 (vinte e cinco) dias do mês de julho de 2017, realizou-se na Sala Laboratório de Estatística 1, no *Campus* de Ji-Paraná, a Sessão de Apresentação e Defesa do Trabalho de Conclusão de Curso (TCC), com o título “**Comparação entre Cadeias de Markov de Primeira e Segunda Ordem para Ocorrência de Períodos Secos e Chuvosos na Região Amazônica.**”, apresentado pelo acadêmico **Páblo Dias Vieira**. Os trabalhos foram instalados às 17:00 horas pelo presidente da Banca Examinadora, aprovada pelo Departamento e constituída por: Prof^ª. **Dr^ª Roziane Sobreira dos Santos** (Universidade Federal de Rondônia - Departamento de Matemática e Estatística), Prof. **Gr. Bruno Soares de Castro** (Universidade Federal de Rondônia - Departamento de Matemática e Estatística) e Prof^º **Dr. Dilson Henrique Ramos Evangelista** (Universidade Federal de Rondônia - Departamento de Matemática e Estatística), sendo este o orientador do acadêmico. A Banca Examinadora, tendo decidido aceitar o Trabalho de Conclusão de Curso, após a apresentação passou à arguição pública do acadêmico. Encerrando os trabalhos de arguição às 17:55 horas. A Banca Examinadora deu parecer final **Aprovado**, com a nota 100 (**cem**), resultado da média aritmética das notas individuais atribuídas pelos membros da Banca Examinadora. Proclamado o resultado final pelo presidente da Banca, foram encerrados os trabalhos. Para constar, lavrou-se a presente Ata, que é assinada pelos membros da Banca Examinadora e pelo acadêmico.

Ji-Paraná, 25 de julho de 2017.

Gr. Bruno Soares de Castro
Prof^ª. Membro da Banca

Roziane sobreira dos santos
Dr^ª. Roziane Sobreira dos Santos
Prof^ª. Membro da Banca

Dr. Dilson Henrique Ramos Evangelista
Prof^º Orientador e Pres. Banca

Páblo Dias Vieira
Páblo Dias Vieira
Acadêmico

Dedico este trabalho a Deus, o provedor da vida, também a minha família em particular minha mãe Maria e meu pai Helio, e aos meus irmãos Hugo e Mônia pelo apoio.

AGRADECIMENTOS

Talvez a parte mais difícil deste trabalho é agrupar todos os que me auxiliaram e apoiaram ao longo deste curso e também antes dele. Mas, em primeiro lugar agradeço muito a Deus por me conceder vida e saúde por todo este tempo.

Meu maior agradecimento vai para meus pais, Maria e Helio, que tiveram contribuição direta e indiretamente em minha formação. Senão fosse por eles nada disto seria possível. Toda a minha família sempre esteve de prontidão em me ajudar no que fosse preciso. Meus irmãos, Mônia e Hugo, que não mediram esforços para me ajudar, mesmo não sabendo qual era o fim e a pessoa que amo, Niely. Obrigado por tudo!

Tenho que prestar meus agradecimentos aos meus tios e tias que sempre reservaram um lugar em seus lares para que eu pudesse me hospedar pelo tempo que precisasse. Obrigado muito tia Créia e família, Tia Iolanda (“Landa”) e Elói, e Tio Natan e família. Também agradeço aos meus outros tios, primos e meus amigos, nos quais me convidaram a ficar em suas casas, mas não tive oportunidade de fazê-los, mesmo assim obrigado.

Agora quero agradecer aos grandes “mestres” (não referindo a titulação) que transmitiram seus valiosos conhecimentos ao longo do curso, obrigado a todos sem exceção. Agradeço a meu estimado orientador professor Dr. Dilson pelos excelentes ensinamentos e lições. Também agradeço a professora Dr.^a Roziane, que foi outra pessoa que transmitiu seus conhecimentos a mim. Quero lembrar também do professor Dr. Ariveltom de minha primeira Iniciação Científica, por ter me dado aquela oportunidade. Estou muito grato por tudo que fizeram.

E, não poderia deixar de agradecer meus amigos de faculdade, em especial o Luis e Tuannhy que trago desde o início do curso, “sobrevivendo” como eu. Obrigado por estarem junto comigo durante este tempo. Tantos outros também que gostaria de falar, mas o espaço é curto, vou resumir assim: muito obrigado aos amigos de Ouro preto, gosto muito de vocês; as meninas de Teixeirópolis e Urupá foi um prazer conhece-las; a meus amigos e amigas de Ji-Paraná foi muito bom contar com vocês; agradeço aos meus companheiros das viagens diárias da Asacam e aos amigos que conheci dos outros cursos da instituição. Obrigado turma.

Muito obrigado a todos e todas, mesmo não tendo citados todo mundo, vocês foram importantes para mim!!!

“Fracassar depois de longa perseverança é muito mais sublime que nunca ter feito um esforço suficientemente bom para ser chamado de fracasso.”

George Eliot (1819 – 1880)

RESUMO

VIEIRA, P. D. **Comparação entre Cadeias de Markov de Primeira e Segunda Ordem para Ocorrência de Períodos Secos e Chuvosos na Região Amazônica.** Monografia (Bacharelado em Estatística) Departamento de Matemática e Estatística, Universidade Federal de Rondônia, 48 p., 2017.

Este trabalho objetivou a comparação com cadeias de Markov de primeira e segunda ordem para as probabilidades de ocorrências de períodos secos e chuvosos em uma região Amazônica, especificamente no Município de Ji-Paraná, Rondônia. A ocorrência de precipitação pluvial pode ser condicionada ao que ocorreu em momentos anteriores. Nesse sentido ao considerar que a ocorrência de chuva é dependente das condições de tempo passadas, esta característica é muito bem aceita na teoria das cadeias de Markov. Estas por sua vez explicam a ocorrência de um evento baseado em eventos anteriores, e a partir daí é possível ligá-las a distribuições de probabilidades para geração de quantidades de precipitação. Sendo assim o estudo foi baseado em dados da precipitação diária no Município de Ji-Paraná entre o período de 2001 a 2015 obtidos no sítio eletrônico da AGRITEMPO. Foram encontradas as estimativas das probabilidades de transição com base nas cadeias de Markov de primeira e segunda ordem e escolhida a ordem apropriada pelos critérios de seleção AIC e BIC. Em sequência foi realizado o ajuste das distribuições Exponencial, Gama e Weibull para as médias mensais do período e testadas a aderência das distribuições pelo teste Kolmogorov-Smirnov a 5% de significância. Os resultados mostraram que a melhor ordem da cadeia de Markov para determinar as probabilidades de ocorrência de chuva em Ji-Paraná é a segunda, na prática isto significa que a partir do conhecimento da ocorrência de dois dias anteriores podemos estimar as probabilidades de ocorrência de chuva no dia atual. Quanto as distribuições de probabilidades, a distribuição Exponencial ajustou-se aos meses de maio a setembro, a Gama e Weibull ajustaram-se igualmente aos meses de agosto a maio. Com base nos resultados encontrados poderão ser realizadas simulações de ocorrência e quantidade de precipitação em séries temporais com poucas observações no Município de Ji-Paraná por meio da cadeia markoviana de ordem dois e com as distribuições de probabilidades correlatadas.

Palavras-Chave: Cadeia de Markov, Distribuições de Probabilidades, Ocorrência de Chuva.

ABSTRACT

VIEIRA, P. D. **Comparison between First and Second Order Markov Chains for Occurrence of Dry and Rainy Periods in the Amazon Region.** Monograph (Bachelor of Statistics) Department of Mathematics and Statistics, Federal University of Rondônia, 48 p., 2017.

This work aimed to compare the first and second order Markov chains for the probabilities of occurrence of dry and rainy periods in an Amazon region, specifically in the municipality of Ji-Paraná, Rondônia. The occurrence of rainfall can be conditioned to what occurred in previous times. In this sense, when considering that the occurrence of rain is dependent on the past conditions of time, this characteristic is very well accepted in the theory of Markov chains. These in turn explain the occurrence of an event based on previous events, and from there it is possible to link them to probability distributions to generate amounts of precipitation. Therefore, the study was based on data from the daily rainfall in the municipality of Ji-Paraná between the period from 2001 to 2015 obtained on the AGRITEMPO website. Transition probability estimates were found based on the first and second order Markov chains and the appropriate order was chosen by the AIC and BIC selection criteria. In the sequence, the Exponential, Gamma and Weibull distributions were adjusted for the monthly averages of the period and the adherence of the distributions by the Kolmogorov-Smirnov test at 5% of significance was tested. The results showed that the best order of the Markov chain to determine the probability of occurrence of rain in Ji-Paraná is the second, in practice this means that from the knowledge of the occurrence of two previous days we can estimate the probability of occurrence of rain On the current day. As for the distributions of probabilities, the Exponential distribution adjusted for the months of May to September, Gama and Weibull also adjusted for the months of August to May. Based on the results, simulations of the occurrence and amount of precipitation can be carried out in time series with few observations in the municipality of Ji-Paraná through the Markovian chain of order two and correlated probabilities distributions.

Keywords: Markov Chain, Probability Distributions, Rain Occurrence.

LISTA DE ILUSTRAÇÕES

Figura 1. Estrutura genérica de um modelo de simulação paramétrico baseado em números pseudo-aleatórios.....	19
Figura 2. Quantidade de dias chuvosos e secos equivalentes ao período de 2001 a 2015 no município de Ji-Paraná, RO	29
Figura 3. Probabilidades de transição estimadas pela cadeia de Markov de 1ª ordem para ocorrência ou não de precipitação no município de Ji-Paraná: P_{00} (dia seco dado anterior seco), P_{10} (dia chuvoso dado anterior seco), P_{01} (dia seco dado anterior chuvoso) e P_{11} (dia chuvoso dado anterior chuvoso).	30
Figura 4. Probabilidades de transição estimadas pela cadeia de Markov de 2ª ordem para a não ocorrência de precipitação no município de Ji-Paraná: P_{000} (dia seco precedido de dois dias secos), P_{100} (dia seco precedido por um dia seco e outro chuvoso), P_{010} (dia seco precedido por um dia chuvoso e outro seco) e P_{110} (dia seco precedido de dois dias chuvosos).	32
Figura 5. Probabilidades de transição estimadas pela cadeia de Markov de 2ª ordem para a ocorrência de precipitação no município de Ji-Paraná: P_{001} (dia chuvoso precedido de dois dias secos), P_{101} (dia chuvoso precedido por um dia seco e outro chuvoso), P_{111} (dia chuvoso precedido por um dia chuvoso e outro seco) e P_{111} (dia chuvoso precedido de dois dias chuvosos)	33

LISTA DE TABELAS

Tabela 1. Arranjo para contagem das transições entre estados em uma cadeia de Markov de 2ª ordem.....	22
Tabela 2. Médias mensais, desvio-padrão, coeficiente de variação, valores extremos mensais de precipitação pluviométrica em Ji-Paraná, RO, para o período de 2001 a 2015.	28
Tabela 3. Estimativas das probabilidades transição da cadeia de Markov de primeira ordem para o Município de Ji-Paraná com base no período de 2001 a 2015.	30
Tabela 4. Estimativas das probabilidades transição da cadeia de Markov de segunda ordem para o Município de Ji-Paraná com base no período de 2001 a 2015.	31
Tabela 5. Parâmetros de ajuste das distribuições de probabilidades Exponencial, Gama e Weibull para as médias de precipitação diárias de cada mês e teste de Kolmogorov-Smirnov para aderência das distribuições estatísticas aos dados.	34
Tabela 6. Resumo das dritribuições de probabilidades com melhor aderência e ordem mais apropriada para a ocorrência de dias secos e chuvosos.	35

ABREVIATURAS E SIGLAS

AGRITEMPO	Sistema de Monitoramento Agroclimatológico
AIC	Critério de Informação Akaike
BIC	Critério de Informação de Bayes
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária
KS	Kolmogorov-Smirnov
mm	Milímetro

LISTA DE EQUAÇÕES

- 1 Probabilidades de transição da cadeia de Markov de 1ª ordem
- 2 Estimativa de probabilidades de transição da cadeia de Markov de 1ª ordem
- 3 Forma genérica de estimação das transições da cadeia de Markov de 1ª ordem
- 4 Probabilidades de transição de uma cadeia de Markov de 2ª ordem
- 5 Forma genérica de estimação das transições da cadeia de Markov de 2ª ordem
- 6 Função de verossimilhança para cadeias de Markov de s estados e ordem um
- 7 Função de verossimilhança para cadeias de Markov de s estados e ordem dois
- 8 Critério de Akaike
- 9 Critério de Bayes
- 10 Função de densidade de probabilidade Exponencial
- 11 Função de densidade de acumulada Exponencial
- 12 Função de densidade de probabilidade Gama
- 13 Estimador do parâmetro γ da distribuição Gama
- 14 Estimador do parâmetro β da distribuição Gama
- 15 Função de densidade de probabilidade Weibull
- 16 Estimador do parâmetro β da distribuição Weibull
- 17 Estimador do parâmetro α da distribuição Weibull
- 18 Módulo do desvio máximo observado

SUMÁRIO

1. INTRODUÇÃO	14
1.1. OBJETIVOS	16
1.1.1. OBJETIVO GERAL	16
1.1.2. OBJETIVOS ESPECÍFICOS	16
2. FUNDAMENTAÇÃO TEÓRICA	17
2.1. REVISÃO DA LITERATURA	17
2.2. DETERMINAÇÃO DA OCORRÊNCIA DE PRECIPITAÇÃO	20
2.2.1. Cadeia de Markov	20
2.2.2. Primeira Ordem	21
2.2.3. Segunda Ordem	22
2.2.4. Escolha da Ordem Ótima	23
2.3. DETERMINAÇÃO DA QUANTIDADE DE PRECIPITAÇÃO	24
2.3.1. Exponencial	24
2.3.2. Gama	24
2.3.3. Weibull	25
2.3.4. Teste de Aderência de Kolmogorov-Smirnov	26
3. METODOLOGIA.....	27
4. RESULTADOS E DISCUSSÕES	28
5. CONCLUSÕES E CONSIDERAÇÕES	36
REFERÊNCIAS	37
APÊNDICE A – ROTINA DE ANÁLISE EM AMBIENTE R	41

1. INTRODUÇÃO

O planejamento de muitas atividades econômicas e processos ambientais são dependentes das condições meteorológicas, principalmente da precipitação pluvial. A maior parte das atividades agrícolas da região Norte do Brasil, por exemplo, são baseadas na precipitação e, o conhecimento de sua variabilidade é de fundamental importância.

A precipitação pluvial é um dos elementos climáticos que exerce maior influência sobre as condições ambientais. Além do seu efeito direto no balanço hídrico, também interfere direta e indiretamente em outros elementos climáticos, como nas temperaturas do ar e do solo, na umidade do ar e que, em conjunto, atuam como fatores básicos no comportamento dos sistemas hidrológicos (ZANNETTI *et al.*, 2006).

Esta variável é encontrada na forma de chuva e se distribui de maneira irregular no tempo e no espaço. Sua presença ou não tem influência no planejamento de algumas atividades humanas: pois, se ocorre de maneira adequada auxilia o abastecimento de água para consumo, favorece as atividades ligadas à pesca, agricultura e pecuária, e beneficia o cultivo e crescimento das plantas, já sua ausência, mesmo que por curtos períodos, pode levar a grandes prejuízos econômicos (MOURA *et al.*, 2015). Assim torna-se necessário compreender tal comportamento meteorológico para avaliar e quantificar adequadamente seus efeitos.

Modelos estocásticos vêm sendo utilizados por diversos pesquisadores para identificação do comportamento da chuva. Tais modelos consideram a precipitação pluvial como um processo aleatório, a partir disto é feita a geração de séries sintéticas da sequência de chuvas com as mesmas características estatísticas da série histórica (NISHIJIMA, 2004; DETZEL, 2009).

Segundo Nishijima (2004) a modelagem estocástica da precipitação pluvial possibilita identificar as características locais de chuva e fornecer as probabilidades da sua ocorrência influenciar as condições futuras da precipitação pluvial, possibilitando a otimização das atividades econômicas e ambientais com base na aplicação destes modelos.

Na maioria dos processos de geração de séries temporais de chuvas, são identificadas duas etapas: a estimação das probabilidades da transição entre dias secos

e chuvosos (ou processo markoviano) e a determinação da distribuição de probabilidades da precipitação pluvial (ANDRADE JÚNIOR, 2001; DETZEL, 2009).

Apesar da ocorrência ou não de precipitação ser decorrente das condições atmosféricas do dia, esta poderá ser condicionada ao tempo passado, ou seja, o que ocorre em um ou mais dias anteriores (AZEVEDO, 1990). Desta forma, a ocorrência de chuvas tende a ter dependência estocástica com o aumento das sequências de dias secos e chuvosos. Esta persistência meteorológica, segundo Azevedo (1990) poderá ser descrita por um modelo de Markov, com ordem apropriada ao grau de dependência condicional.

Após definido o dia chuvoso, por um modelo de ocorrência, pode-se adotar um modelo para geração da quantidade de precipitação pluvial baseados em distribuições de probabilidades. A exemplo desses modelos temos a distribuição Gama de dois parâmetros (COE e STERN, 1982) e a distribuição Weibull (ZUCCHINI e ADAMSON, 1984). A forma mais comum de se representar a distribuição da quantidade de precipitação em modelos climáticos estocásticos é pela distribuição Gama (WILKS, 2006).

Neto *et al.* (2005) comparou as distribuições de probabilidades Gama, Exponencial, Weibull, Normal e Log-Normal, para o ajuste de suas funções de densidade de probabilidade para dados diários de precipitação pluvial em Piracicaba, São Paulo. Verificou-se que a distribuição Gama obteve melhor desempenho, e em seguida a distribuição Weibull, com desempenho próximo da distribuição Gama, a mais precisa para valores médios diários no município entre as distribuições comparadas.

Porém as distribuições Gama e Weibull tem limitações quando se tratam de períodos de seca, pois em ambas não é possível ajustar seus parâmetros quando há valores nulos. Nesse sentido uma alternativa para os períodos secos é a distribuição Exponencial. O modelo exponencial foi utilizado para descrição do período de seca, em Shiau e Morrades (2009), para dados de estações no norte do Iran. Foi indicada a adequação do modelo exponencial para o período de seca, sendo que a qualidade foi observada por meio de gráficos. Catalunha (2000) também verificou que para períodos secos a distribuição Exponencial teve melhor aderência para precipitação no Estado de Minas Gerais.

Paiva e Clarke (1997) modelaram a precipitação da região amazônica por meio das cadeias de Markov de 1ª ordem. O estudo ressaltou que seria importante a aplicação da modelagem por cadeias de Markov de ordem superior a apresentada. Com isso modelar a ocorrência de precipitação com uma cadeia Markoviana de segunda ordem pode ser uma alternativa para compreensão mais precisa do clima amazônico, visto que o modelo será condicional a dois períodos de tempos passados para estabelecer as probabilidades de chuva no tempo atual. Nesse sentido alguns autores que Almeida (1995) cita relatam que há “alguns casos [...] em que a cadeia de Markov de 1ª ordem não é a mais adequada para descrever a ocorrência de precipitação diária, havendo necessidade de se considerarem ordens de dependências maiores que um, nos modelos estocásticos”.

1.1. OBJETIVOS

1.1.1. OBJETIVO GERAL

Portanto, este estudo objetiva realizar comparação com cadeias de Markov de primeira e segunda ordem para as probabilidades de ocorrências de períodos secos e chuvosos em uma região Amazônica.

1.1.2. OBJETIVOS ESPECÍFICOS

- a)** Estimar as probabilidades de transição entre períodos secos e chuvosos por meio de cadeia de Markov de primeira ordem e segunda ordem;
- b)** Encontrar os parâmetros das distribuições de probabilidades Exponencial, Gama e Weibull para determinação das quantidades precipitação pluvial;
- c)** Definir pelo teste de Kolmogorov-Smirnov a distribuição de probabilidades que melhor se adere a precipitação mensal;
- d)** Escolher a ordem apropriada da cadeia de Markov pelos critérios de avaliação AIC e BIC.

2. FUNDAMENTAÇÃO TEÓRICA

2.1. REVISÃO DA LITERATURA

Em meio a grande demanda de informação pluviométrica presente no país, o uso de técnicas estatísticas para preenchimento de dados faltantes vem se tornando de grande importância, visto que, estas técnicas permitem aproximar o padrão de chuvas de uma região.

Dados pluviométricos disponíveis e confiáveis são fundamentais para uma eficaz previsão de eventos impactantes ao meio ambiente. Para contornar o problema de inexistência ou inequabilidade de dados é comum gerar uma série sintética produzida por modelo estocástico. Segundo Wilks *apud* ZANETTI *et al.* (2006), série sintética é o resultado da associação de séries de dados reais com números aleatórios produzidos por algoritmos computacionais, a fim de gerar sequências de números aleatórios que se assemelham aos dados climáticos reais. No Brasil, a aplicação desses procedimentos é de grande importância visto que as extensões das séries históricas disponíveis normalmente são reduzidas, e a exemplo disso temos os dados pluviográficos (ZANETTI *et al.*, 2006).

Métodos estocásticos para preenchimento de dados faltantes, conforme Furlan (2009) são importantes para utilização de modelos em culturas agrícolas e, também para a racionalização dos recursos hídricos destinados à irrigação.

Atualmente, o número de pesquisas que fazem a utilização de técnicas estocásticas para a geração de dados sintéticos aumentou. Esses métodos são caracterizados não só pela ocorrência ou não de precipitação num determinado dia que depende somente das condições atmosféricas reinantes naquele dia e local, mas está condicionada também ao tempo passado, ou seja, a situação ocorrida no dia ou dias anteriores (AZEVEDO e LEITÃO, p. 390, 1990). Neste contexto um modelo de Markov poderá descrever apropriadamente o padrão de dias secos e chuvosos em uma dada região.

No âmbito das cadeias de Markov de ordem um, trabalhos como de Back *et al.* (2014) utilizaram o método para simular as sequências de dias secos e chuvosos no município de Saudades, Santa Catarina; Andrade Júnior *et al.* (2001) fizeram a simulação da precipitação diária em Parnaíba e Teresina, Piauí, por meio da cadeia de Markov de

1ª ordem; Andrade *et al.* (2009) aplicaram as probabilidades condicionais junto com o processo de Markov para o ocorrência de dias secos e chuvosos no município de Garanhuns, Pernambuco. Em todos os trabalhos acima citados, o método manteve as características das séries observadas pelas séries simuladas. Além dos autores acima, outros trabalhos utilizaram o método tanto no Brasil como fora. Entre eles temos Bezerra *et al.* (2015), Machado *et al.* (2014), Baú *et al.* (2013), Barkotulla (2010), Nishijima (2004), Viana *et al.* (2002), Sediya *et al.* (1996), e outros.

Estudos realizando estimativas com cadeias de Markovianas de segunda ordem no Brasil são raríssimos de encontrar. Mas dentre estes Rocha Filho e Pinto (2015) trabalharam com geração estocástica de precipitação para cadeias de Markov de segunda ordem em três localidades do Estado de Minas Gerais. Já Keller Filho (2006) trabalhou com o método de Markov de até terceira ordem na modelagem de dias secos e chuvosos na região Centro-Oeste do País. Basak (2014) fez diferentes modelos de Markov para ocorrência de chuvas monozônicas em diferentes zonas de West Bengal, Índia, entre eles o de segunda ordem.

Punyawardena e Kulasiri (1996) compararam cadeias de Markov de primeira e segunda ordem para precipitação na Zona Seca de Sri Lanka. O estudo não encontrou diferença significativa entre a série simulada e os registros históricos do local de estudo para ambas as ordens. Mas o desempenho global do modelo de ordem dois sobressaiu ao modelo de ordem um, visto que, o mesmo pareceu ser mais adequado para aplicações agrícolas assim como para estudar e projetar a climatologia da precipitação da Zona Seca de Sri Lanka.

Furlan (2009) desenvolveu, no Estado de Rondônia, um estudo para verificar se os índices de desmatamento no Estado eram suficientes para contribuir para a modificação do regime hídrico e pluviométrico. Ao decorrer do trabalho utilizou-se o método da cadeia de Markov de primeira ordem para simular uma série de dados a fim de preencher falhas e cumprir o objetivo estabelecido. Após a análise, chegou-se à conclusão que o método poderia ser utilizado no Estado de Rondônia para preencher dados pluviométricos faltantes. Porém, em algumas localidades foram encontradas inconsistências entre os dados simulados e a série real, mas no geral foi bem aceita e gerou resultados satisfatórios.

Estudos como Andrade Júnior *et al.* (2001) e Rocha Filho *et al.* (2009) são baseadas em duas etapas para construção de modelo de forma paramétrica, na qual as ocorrências são determinadas e os dias chuvosos associados a uma distribuição estatística para a obtenção das alturas precipitadas. Essas etapas são fundamentais visto que a precipitação pluvial diária é considerada como sendo $Z_t = X_t Y_t$ onde Z_t representa a precipitação pluvial diária em mm, X_t a ocorrência e Y_t a quantidade de chuva como descrito por Assis e Villa Nova (1994).

A Figura 1 mostra o processo de geração de precipitação com base em modelo paramétrico, onde os processos estocásticos e as distribuições estatísticas utilizam-se de números pseudoaleatórios, tanto para determinação da ocorrência como das alturas de precipitação.

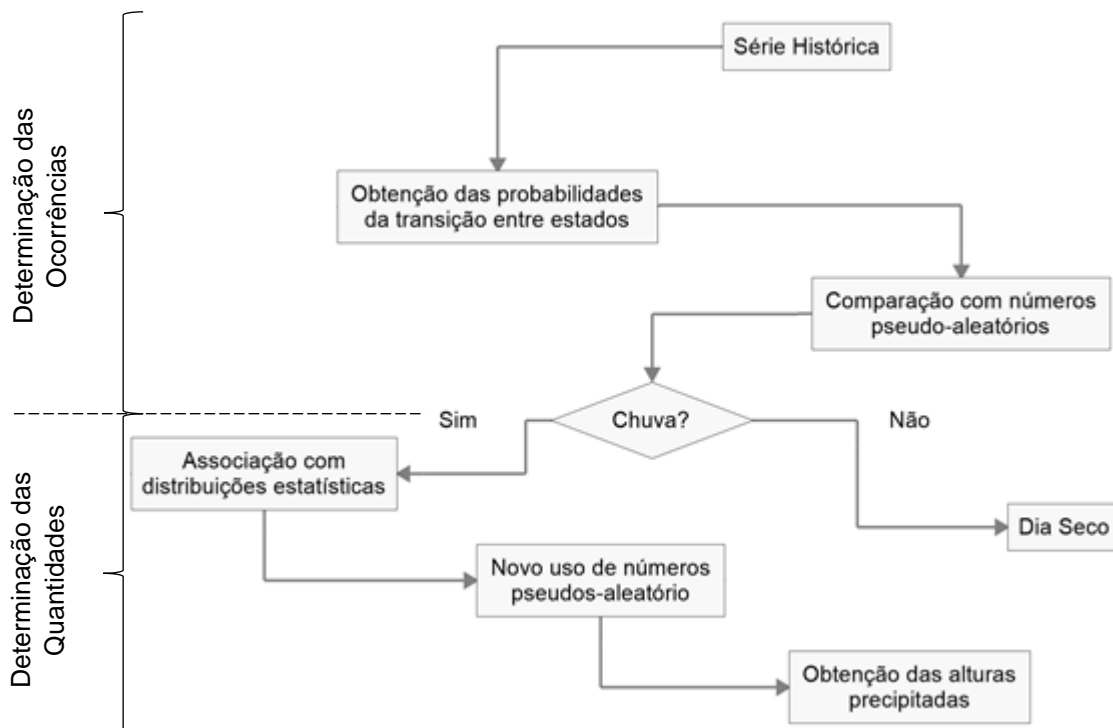


Figura 1. Estrutura genérica de um modelo de simulação paramétrico baseado em números pseudo-aleatórios

FONTE: Detzel (2009)

Portanto, é de suma importância a determinação da ocorrência de dias secos e chuvosos para que sejam obtidas as quantidades de precipitação com a utilização de distribuições estatísticas de probabilidades. Com isso, este texto vai dedicar duas seções para que sejam abordados mesmo que sucintamente estes dois temas.

2.2. DETERMINAÇÃO DA OCORRÊNCIA DE PRECIPITAÇÃO

A determinação de dias chuvosos ou secos, em muitos estudos, está relacionada com uma probabilidade de ocorrência. Quando problemas envolvem probabilidades, estes são atrelados ao conceito de probabilidade, onde os resultados tornam-se aleatórios e independentes do resultado anterior. No caso da precipitação em escala diária, dificilmente uma ocorrência atual será independente de uma ocorrência anterior, mesmo com os conceitos de probabilidade.

Detzel (2009) explica que não se pode obter uma generalização do grau de dependência entre os eventos devido à complexidade das dinâmicas atmosféricas, mas que mesmo assim, tendo como base a constatação das tendências analisadas ao longo dos anos em diversos lugares do planeta, a aplicação de processos estocásticos aparece como sendo uma boa ferramenta para a determinação das ocorrências desejadas.

Coe e Stern (1982) recomendam o uso da cadeia de Markov de primeira ou de segunda ordem para descrever a ocorrência diária de chuva, na maioria dos casos. Esses autores recomendam a cadeia de Markov de primeira ordem, que descreve satisfatoriamente a ocorrência de chuva, devido a maior simplicidade no cálculo de outras quantidades e pelo menor número de parâmetros a serem estimados.

2.2.1. Cadeia de Markov

Os modelos de cadeia de Markov podem ser descritos por duas propriedades. A primeira é em relação ao número de valores diferentes que a variável pode ter (conhecido como o "estado") e a segunda ao número de valores anteriores usados para determinar as probabilidades de transição de estado para estado (conhecida como "ordem"). No caso de ocorrência de precipitação, a variável X_k é binária, tendo apenas dois estados: ocorrência ($X_k = 1$) ou não-ocorrência ($X_k = 0$) de precipitação (SCHOOFF *et al.*, 2008).

Sendo assim a variável binária para dias secos e úmidos pode ser representada como:

$$X_k = \begin{cases} 0, & \text{para não ocorrência de precipitação no dia } k \\ 1, & \text{para ocorrência de precipitação no dia } k \end{cases}$$

Onde, $k = 1, 2, 3, \dots$, etc.

A determinação das mudanças entre estados é de fundamental importância para a construção de todo o processo. O cálculo dessas probabilidades se dá através da

contagem dos elementos presentes nos registros históricos da localidade desejada (k). Sendo assim, para este estudo vamos conhecer dois tipos de probabilidades de transição da cadeia de Markov: primeira e segunda ordem.

2.2.2. Primeira Ordem

Para a cadeia de Markov de primeira ordem, as probabilidades de transição que controlam o estado futuro da variável estudada dependem apenas do seu estado atual. Sabendo que no tempo k a variável X está no estado 0 (não ocorre precipitação e $X_k = 0$), ou no estado 1 (ocorre precipitação e $X_k = 1$), as probabilidades de transição condicional são computadas no tempo k+1, onde o tempo passado é de um dia (WILKS, p. 342, 2006). Logo,

$$\begin{cases} p_{00} = \Pr(X_{k+1} = 0 | X_k = 0) \\ p_{01} = \Pr(X_{k+1} = 1 | X_k = 0) \\ p_{10} = \Pr(X_{k+1} = 0 | X_k = 1) \\ p_{11} = \Pr(X_{k+1} = 1 | X_k = 1) \end{cases} \quad (1)$$

O procedimento de estimativa de probabilidades de transição consiste em calcular as frequências relativas condicionais, da seguinte forma:

$$p_{ij} = \begin{cases} p_{00} = \frac{n_{00}}{n_{00} + n_{01}} = \frac{n_{00}}{n_0} \\ p_{01} = \frac{n_{01}}{n_{00} + n_{01}} = \frac{n_{01}}{n_0} = 1 - p_{00} \\ p_{11} = \frac{n_{11}}{n_{10} + n_{11}} = \frac{n_{11}}{n_1} \\ p_{10} = \frac{n_{10}}{n_{10} + n_{11}} = \frac{n_{10}}{n_1} = 1 - p_{11} \end{cases} \quad (2)$$

ou de forma genérica,

$$\hat{p}_{01} = \frac{n_{01}}{n_{0*}} \quad e \quad \hat{p}_{11} = \frac{n_{11}}{n_{1*}} \quad (3)$$

Em que, n_{01} é o número de transições do estado 0 para o estado 1, n_{11} é o número de pares de etapas de tempo em que o sistema mantém no estado 1, n_{0*} é o número de estados 0 na série seguido por outro ponto de dados e n_{1*} é o número de estados 1 na

série seguido por outro ponto de dados. Isto é, $n_{1*} = n_{10} + n_{11}$ e $n_{0*} = n_{00} + n_{01}$ (WILKS, p. 342, 2006).

2.2.3. Segunda Ordem

A dependência de tempo de segunda ordem significa que as probabilidades de transição dependerão dos estados em defasagens de um e dois períodos de tempo. Então, as probabilidades de transição para uma cadeia de Markov de segunda ordem requerem três indicadores: o primeiro refere-se ao estado no instante $k-1$, o segundo ao estado no instante k , e o terceiro especifica o estado no instante $k+1$ (WILKS, p. 342, 2006). A notação para as probabilidades de transição de uma cadeia de Markov de segunda ordem pode ser definida como:

$$p_{hij} = \Pr(X_{k+1} = j \mid X_k = i, X_{k-1} = h) \quad h, i, j = 0, 1 \quad (4)$$

As estimativas das probabilidades de transição para obtenção da cadeia de Markov de segunda ordem é dada pela fórmula:

$$\hat{p}_{hij} = \frac{n_{hij}}{n_{hi*}} \quad (5)$$

Isto é, sabendo que o valor da série temporal no tempo $k-1$ era $X_{k-1}=h$ e no tempo k , $X_k=i$, a probabilidade de que o estado futuro da série temporal $X_{k+1}=j$ seja p_{hij} . Então para que possam ser calculadas as estimativas das probabilidades de transição da cadeia de Markov de segunda ordem são necessárias as contagens das transições entre estados, como esquematizado na Tabela 1.

Tabela 1. Arranjo para contagem das transições entre estados em uma cadeia de Markov de 2ª ordem.

X_{t-1}	X_t	$X_{t+1} = 0$	$X_{t+1} = 1$	Marginais Totais
0	0	n_{000}	n_{001}	$n_{00*} = n_{000} + n_{001}$
0	1	n_{010}	n_{011}	$n_{01*} = n_{010} + n_{011}$
1	0	n_{100}	n_{101}	$n_{10*} = n_{100} + n_{101}$
1	1	n_{110}	n_{111}	$n_{11*} = n_{110} + n_{111}$

FONTE: Adaptado de Wilks (2006, p 350)

Uma vez calculados todos os parâmetros das cadeias de Markov, poderão ser realizadas os testes para escolha da melhor ordem.

2.2.4. Escolha da Ordem Ótima

A escolha da ordem mais apropriada da cadeia de Markov para descrever a ocorrência de precipitação diária deve-se basear em uma série de dados, levando-se em conta procedimentos do teste de hipótese, do teste chi-quadrado ou, ainda, do princípio da máxima verossimilhança (ALMEIDA, 1995).

Com isso a determinação da ordem do modelo markoviano mais apropriada para a certo período pode-se utilizar de alguns critérios de avaliação. A exemplo temos os critérios de AIC (AKAIKE, 1974) e BIC (SCHWARZ, 1978), os quais visam identificar o melhor modelo com o menor número de parâmetros e baseiam-se nas funções de log-verossimilhança das probabilidades de transição, dadas por:

$$L_1 = \sum_{i=0}^{s-1} \sum_{j=0}^{s-1} N_{ij}(k) \ln [p_{ij}(k)] \quad (6)$$

$$L_2 = \sum_{h=0}^{s-1} \sum_{i=0}^{s-1} \sum_{j=0}^{s-1} N_{hij}(k) \ln [p_{hij}(k)] \quad (7)$$

onde N representa o número de dias presentes nas séries históricas com as transições i, j ou h e p representa as probabilidades de transição.

Os critérios são baseados no princípio da parcimônia. A ordem ótima do modelo é obtida a partir de uma equação que releva a aderência do ajuste (através das funções de verossimilhança) e uma penalidade que aumenta proporcionalmente com o número de parâmetros a ser utilizado. Para o Critério de Akaike (AIC) (AKAIKE, 1974) e o Critério de Bayes (BIC) (SCHWARZ, 1978), as equações são dadas, respectivamente, por:

$$AIC(m) = -2L_m + 2s^m(s-1) \quad (8)$$

e

$$BIC(m) = -2L_m + s^m(\ln n) \quad (9)$$

Para ambas as equações, m representa a ordem da cadeia de Markov a ser testada, s o número de estados e n o tamanho da amostra.

2.3. DETERMINAÇÃO DA QUANTIDADE DE PRECIPITAÇÃO

O ajuste de modelos probabilísticos a dados diários de chuva é capaz resumir sucintamente estes dados, além de ser uma técnica eficiente para a análise dessas informações (STERN & COE, 1982 *apud* ALMEIDA, 1995). Conforme Catalunha *et al.* (2002), cada distribuição de frequência apresenta uma certa forma, que pode ser aproximada através da utilização de equações de densidade probabilidade com alguns parâmetros extraídos da amostra em questão. Portanto, a utilização ou não de uma distribuição de probabilidades consiste na capacidade desta em estimar, por meio seus parâmetros, os valores dos dados observados, sendo que esta capacidade é medida através de testes de aderência.

2.3.1. Exponencial

A distribuição exponencial é uma das matematicamente mais simples, e geralmente é aplicada para das com forte assimetria, ou seja, apresentando uma função de “J” invertido. Esta distribuição é caracterizada por sua função de taxa de falha constante, sendo que possui inúmeras aplicação em várias áreas do conhecimento, inclusive em variáveis hidrológicas (CATALUNHA, 2002; PAULA JÚNIOR, 2016). A função de densidade de probabilidade é dada por:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{para } x \geq 0 \\ 0 & \text{para } x < 0 \end{cases} \quad (10)$$

sua função acumulada é do tipo:

$$F(x) = \int_0^{\infty} f(x) = 1 - e^{-\lambda x} \quad (11)$$

Sendo que o único parâmetro da distribuição (λ) visto é estimado pelo inverso da média.

2.3.2. Gama

A distribuição Gama é a mais utilizada para ajuste de totais de chuvas de períodos curtos com duração de até 30 dias (ASSIS *et al.*, 1996). Sua função de probabilidades é dada por:

$$f(X) = \frac{1}{\Gamma(\gamma)\beta^\gamma} X^{\gamma-1} e^{-\frac{X}{\beta}} \quad (12)$$

Os parâmetros α e β que possibilitam o cálculo da Distribuição Gama incompleta para uma dada variável aleatória foram estimados pelo método da máxima verossimilhança (ASSIS *et al.*, 1996), conforme as equações:

$$\gamma = \frac{1}{4A} \left(1 + \sqrt{1 + \frac{4A}{3}} \right) \quad (13)$$

$$\beta = \frac{\bar{X}}{\gamma} \quad (14)$$

sendo,

$$A = \ln \bar{X} - \frac{1}{n} \sum_{k=1}^n \ln(X_k)$$

onde \bar{X} é a precipitação mensal média, \ln o logaritmo neperiano, n o número de dados.

2.3.3. Weibull

A distribuição Weibull, conforme Catalunha *et al.* (2002) a distribuição de probabilidades tem função de densidade de probabilidade dada, na forma mais comum, por:

$$f(x) = \frac{\gamma}{\beta} \left(\frac{x - \alpha}{\beta} \right)^{\gamma-1} e^{-\left(\frac{x - \alpha}{\beta} \right)^\gamma}; \text{ para } x \geq \alpha \quad (15)$$

onde $F(x) = 0$ para outros intervalos, e $\alpha \geq 0$, $\beta > 0$, $\gamma > 0$ são os parâmetros da distribuição. As estimativas de máxima verossimilhança de α e β são (RODRIGUES, 2013):

$$\hat{\beta} = \left(\frac{1}{n} \sum_{i=0}^n x_i^{\hat{\alpha}} \right)^{1/\hat{\alpha}} \quad (16)$$

$$\hat{\alpha} = \left[\left(\sum_{i=0}^n x_i^{\hat{\alpha}} \ln x_i \right) \left(\sum_{i=0}^n x_i^{\hat{\alpha}} \right)^{-1} - \frac{1}{n} \sum_{i=0}^n \ln x_i \right]^{-1} \quad (17)$$

em que x_i é a precipitação para o dia i , e n o número de dados.

2.3.4. Teste de Aderência de Kolmogorov-Smirnov

O teste não-paramétrico de Kolmogorov-Smirnov avalia a concordância entre a distribuição observada da amostra e uma determinada distribuição teórica. Avalia se os dados amostrais se aproximam razoavelmente de uma determinada distribuição. Para isso utilizamos a função distribuição acumulada observada, compara-se com a teórica, determina-se o ponto em que estas distribuições mais divergem, e testamos se essa divergência é aleatória ou não (CAMPOS, 1976; RODRIGUES, 2013).

Para F , função de distribuição (desconhecida) da população e F_0 , função de distribuição proposta, contínua e completamente especificada, as hipóteses testadas serão:

$$\begin{aligned} H_0: F(X) &= F_0(X), \quad -\infty < X < +\infty \\ H_1: F(X) &\neq F_0(X) \end{aligned}$$

ou de maneira mais clara,

H_0 : O comportamento probabilístico da variável aleatória, em questão pode ser modelado pela seguinte distribuição x de probabilidade.

H_1 : O comportamento probabilístico da variável aleatória, em questão não pode ser modelado pela seguinte distribuição x de probabilidade.

A estatística do teste de Kolmogorov-Smirnov, ou seja, o maior valor das diferenças entre $F_0(X)$ e $F_n(X)$ é dado por:

$$D = \max | F_0(X) - F_n(X) | \quad (18)$$

onde $F_0(X)$ é uma distribuição teórica acumulada e $F_n(X)$ uma distribuição observada em uma amostra de n observações (distribuição empírica).

A regra de decisão é rejeitar H_0 se as estatísticas de teste D é maior que o quantil $1-\alpha$ para o teste Kolmogorov-Smirnov. Caso contrário, não se rejeita a hipótese H_0 .

3. METODOLOGIA

Foram utilizados dados da precipitação pluvial diária, observada em uma estação meteorológica de Ji-Paraná, nas coordenadas -10.75 (Latitude) e -65.75 (Longitude). Os dados abrangem um período de 15 anos (2001 a 2015) e foram obtidos no banco de dados do Sistema de Monitoramento Agroclimatológico (Agritempo), desenvolvido pela Embrapa Informática Agropecuária e pela Unicamp.

A modelagem das probabilidades de ocorrência de precipitação será feita a partir das cadeias de Markov de primeira e segunda ordem e com dois estados (seco e chuvoso). As probabilidades de transição entre os estados seco/chuvoso serão determinadas por séries mensais, considerando as sazonalidades para ajustar os parâmetros do modelo. Será considerado como dia onde não houve ocorrência de precipitação (ou dia seco) aquele que corresponde a uma precipitação pluvial inferior a 1 mm (KELLER FILHO, 2006).

O ajuste para quantidade de precipitação diária foi feito utilizando as distribuições de probabilidades Exponencial, Gama e Weibull. Para a distribuição Exponencial foi ajustado o parâmetro λ . Na distribuição Gama foram ajustados os parâmetros γ e β , respectivamente, parâmetro de forma (adimensional) e parâmetro de escala. Para a distribuição Weibull α é o parâmetro de posição, β o parâmetro que controla as dimensões que a curva assume e γ é o parâmetro de forma (CATALUNHA, 2000).

Para verificar a aderência das distribuições descritas, foi utilizado o teste estatístico de Kolmogorov-Smirnov, ao nível de 5% de significância, que além de não depender do número de classes do agrupamento dos dados, oferece mais vantagens computacionais. E para a escolha da melhor ordem da cadeia de Markov para cada mês foram utilizados os critérios AIC e BIC.

Os dados foram organizados e analisados em conjunto de planilha eletrônica com o software livre R (R CORE TEAM, 2017). O ajuste das distribuições de probabilidades foi feito pelo pacote *fitdistrplus* (DELIGNETTE-MULLER e DUTANG, 2015), para geração das cadeias de Markov foram elaborados linhas de código com funções específicas.

4. RESULTADOS E DISCUSSÕES

Os resultados iniciais do trabalho são referentes a caracterização da variável precipitação pluvial do Município de Ji-Paraná. Neste sentido foi elaborada a Tabela 2 com as médias de chuva, desvio padrão, valor máximo de precipitação do período e coeficiente de variação para acompanhar o volume de dispersão da média de precipitação.

Tabela 2. Médias mensais, desvio-padrão, coeficiente de variação, valores extremos mensais de precipitação pluviométrica em Ji-Paraná, RO, para o período de 2001 a 2015.

Meses	Média (mm)	Desvio Padrão (mm)	Máximo (mm)	Coeficiente de Variação (%)
Jan.	10,86	13,24	106,08	121,9
Fev.	13,49	15,78	82,29	117,0
Mar.	10,31	14,26	108,54	138,2
Abr.	6,82	11,70	74,88	171,5
Mai.	2,29	5,86	47,79	256,0
Jun.	0,90	5,13	66,18	569,7
Jul.	0,43	2,72	36,66	629,7
Ago.	1,05	5,71	80,52	546,0
Set.	2,02	5,67	39,57	280,8
Out.	5,54	10,29	75,06	185,6
Nov.	6,77	10,11	70,11	149,3
Dez.	9,26	11,97	72,39	129,2

FONTE: Elaboração através do *software* R

Foram constatadas que as maiores médias de precipitação ficam entre os meses de dezembro a março, chegando a 13,49 mm a média de precipitação para um dia no mês de fevereiro. Estes meses ainda variam entre 11,97 e 15,78 mm de chuva, enquanto que nos meses de maio a setembro essa quantidade variou menos de 6 mm de chuva, onde as médias de precipitação foram inferiores a 2,29 mm.

Quanto aos valores máximos de chuva para o período de 2001 a 2015 em Ji-Paraná, no mês de março foi encontrado a maior quantidade de precipitação neste período, sendo que este foi de 108,54 mm no dia 24 de março de 2009. O mês de julho é o que teve a menor precipitação máxima, sendo que ainda este mês tem coeficiente de variação de 629,7%, que indica existência de grande dispersão em relação à média do mês.

Na Figura 2 são representados o número de dias secos e chuvosos mensais relativos aos 15 anos de dados da série temporal em estudo. Nesse sentido, a quantidade

de dias com tempo seco foi maior que o número de dias com chuva nos meses de maio, junho, julho e agosto e setembro. A situação é diferente de outubro a abril, onde apresentam maior número de dias com tempo chuvoso no Município de Ji-Paraná, visto que o mês de abril é caracterizado pelo declínio na quantidade de dias com chuva e outubro como o fim da estiagem.

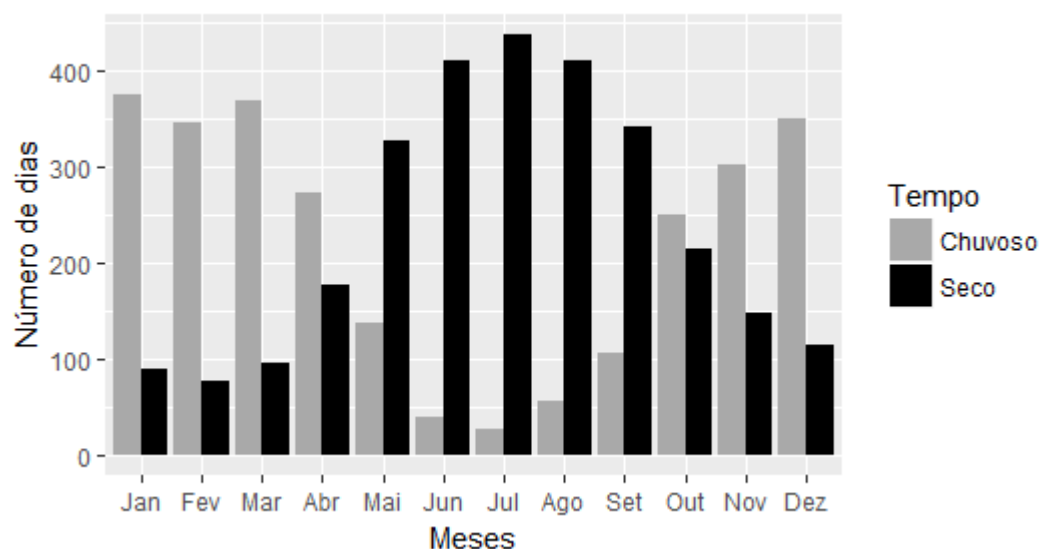


Figura 2. Quantidade de dias chuvosos e secos equivalentes ao período de 2001 a 2015 no município de Ji-Paraná, RO

FONTE: Elaboração em ambiente R.

As informações acima corroboram o que diz Franca (2015), pois no clima da região amazônica existe um período bem definido de estiagem entre os meses de junho a agosto, que é influenciada pela forte subsidência atmosférica exercida pela atuação de anticiclones sobre o continente, inibe a formação de nuvens de chuva no Brasil central, e outro período bem definido de chuvas entre os meses de dezembro a fevereiro, que ocorre em função do aquecimento do continente sul-americano, queda da pressão atmosférica e aumento da convecção tropical.

A Figura 3 mostra a evolução das probabilidades condicionais de transição de dias secos e chuvosos por meses do ano, conforme a primeira ordem a cadeia de Markov. Nota-se que nos três primeiros e nos dois últimos meses do ano, tanto as probabilidades de ocorrer dias secos consecutivos (P_{00}) e ocorrer um dia chuvoso seguido por um dia seco (P_{10}) aparentam ser menores. Percebe-se também que probabilidades condicionais de chuva e de estiagem são inversamente proporcionais em todos os períodos do ano, o que é comum visto que as probabilidades de transição são complementares.

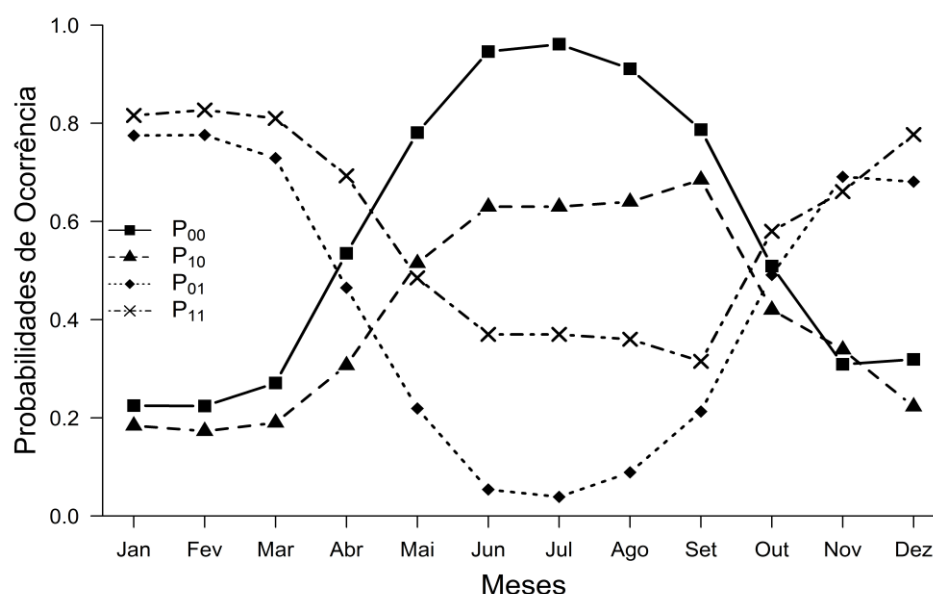


Figura 3. Probabilidades de transição estimadas pela cadeia de Markov de 1ª ordem para ocorrência ou não de precipitação no município de Ji-Paraná: P_{00} (dia seco dado anterior seco), P_{10} (dia chuvoso dado anterior seco), P_{01} (dia seco dado anterior chuvoso) e P_{11} (dia chuvoso dado anterior chuvoso).

FONTE: elaborado em ambiente R

Pela Figura 3 observamos ainda que as probabilidades de transição se aproximam uma das outras no mês de outubro, onde o número de dias chuvosos foi próximo no período correspondente (Figura 2). Nesse sentido, os valores estimados das probabilidades de transição entre dias secos e chuvosos para Ji-Paraná são apresentados na Tabela 3, onde também contém as probabilidades de cada mês de um determinado dia ser chuvoso ou seco além dos critérios AIC e BIC.

Tabela 3. Estimativas das probabilidades transição da cadeia de Markov de primeira ordem para o Município de Ji-Paraná com base no período de 2001 a 2015.

Meses	P_0	P_1	Parâmetros da cadeia de Markov				AIC	BIC
			P_{00}	P_{10}	P_{01}	P_{11}		
Jan.	0,211	0,789	0,225	0,184	0,775	0,816	457,301	465,585
Fev.	0,208	0,792	0,224	0,173	0,776	0,827	404,36	412,455
Mar.	0,241	0,759	0,271	0,19	0,729	0,81	474,658	482,942
Abr.	0,424	0,576	0,535	0,307	0,465	0,693	584,228	592,446
Mai.	0,692	0,308	0,781	0,515	0,219	0,485	538,159	546,443
Jun.	0,918	0,082	0,946	0,63	0,054	0,37	235,439	243,657
Jul.	0,925	0,075	0,961	0,63	0,039	0,37	183,392	191,676
Ago.	0,914	0,086	0,911	0,64	0,089	0,36	318,825	327,109
Set.	0,713	0,287	0,787	0,685	0,213	0,315	493,195	501,413
Out.	0,462	0,538	0,509	0,42	0,491	0,58	642,268	650,552
Nov.	0,367	0,633	0,309	0,339	0,691	0,661	573,637	581,856
Dez.	0,269	0,731	0,319	0,223	0,681	0,777	520,097	528,382

0 = seco, 1 = chuvoso

FONTE: Elaboração através do software R

Como citado anteriormente, outubro apresentou probabilidade de dois dias consecutivos sem chuva (P_{00}) de 0,509 e para dias consecutivos com chuva (P_{11}) de 0,58. Quando tratamos das probabilidades de ocorrência de precipitação precedida por dia onde não houve precipitação (P_{01}), observa-se variação em função do período analisado, desde 0,681 a 0,776 na estação chuvosa e de 0,039 a 0,089 na estação seca. Isto mostra a tendência natural de menores ocorrências de chuva na estação seca e maiores na estação chuvosa em função do dia passado ser seco.

Acima foram descritas as probabilidades de dependência para um dia sobre as condições de precipitação de Ji-Paraná, agora serão apresentadas as probabilidades de transição da cadeia de Markov de segunda ordem e dois estados, seco e chuvoso (Tabela 4), ou seja, dois dias de dependência para determinar a probabilidade de ocorrência do dia atual.

Tabela 4. Estimativas das probabilidades transição da cadeia de Markov de segunda ordem para o Município de Ji-Paraná com base no período de 2001 a 2015.

Meses	Parâmetros da cadeia de Markov								AIC	BIC
	P_{000}	P_{100}	P_{010}	P_{110}	P_{001}	P_{101}	P_{011}	P_{111}		
Jan.	0,333	0,191	0,138	0,193	0,667	0,809	0,862	0,807	418,394	434,962
Fev.	0,267	0,213	0,161	0,175	0,733	0,787	0,839	0,825	366,794	382,983
Mar.	0,37	0,232	0,214	0,184	0,63	0,768	0,786	0,816	422,017	438,585
Abr.	0,674	0,393	0,275	0,32	0,326	0,607	0,725	0,68	485,748	502,185
Mai.	0,812	0,671	0,565	0,463	0,188	0,329	0,435	0,537	431,157	447,725
Jun.	0,963	0,724	0,63	0,632	0,037	0,276	0,37	0,368	181,825	198,262
Jul.	0,964	0,889	0,471	0,9	0,036	0,111	0,529	0,1	159,755	176,323
Ago.	0,922	0,767	0,657	0,6	0,078	0,233	0,343	0,4	270,448	287,016
Set.	0,8	0,736	0,667	0,722	0,2	0,264	0,333	0,278	392,158	408,595
Out.	0,513	0,505	0,454	0,393	0,487	0,495	0,546	0,607	500,91	517,478
Nov.	0,396	0,267	0,382	0,317	0,604	0,733	0,618	0,683	467,133	483,57
Dez.	0,324	0,316	0,232	0,221	0,676	0,684	0,768	0,779	449,747	466,315

0 = seco, 1 = chuvoso

FONTE: Elaboração através do software R

As menores estimativas de transição foram para um dia de chuva precedido de dois secos (P_{001}) entre os meses de junho a agosto. Além disso as probabilidades de 3 dias consecutivos sem chuva (P_{000}) nos meses de estiagem em Ji-Paraná são grandes. No período de chuvas no Município as probabilidades do dia atual chuvoso permaneceram altas, mesmo com dias anteriores sem chuva ou alternando entre dias secos e chuvosos. Já para no período de estiagem estas probabilidades se mantiveram abaixo de 0,529.

Nas Figura 4 são traçadas as probabilidades de transição de dia seco com todas as possibilidades para dois dias anteriores. Nela é possível notar as probabilidades de três dias consecutivos sem chuva (P_{000}) são relativamente maiores em comparação com as demais probabilidades, onde nos meses de estiagem em Ji-Paraná são vistas as maiores probabilidades e em fevereiro esta probabilidade foi a menor. Nesse sentido, nos meses de chuva tres dias consecutivos sem chuva são baixas as chances de acontecerem, enquanto que na época de estiagem tem grande chance de ter três dias consecutivos sem chuva.

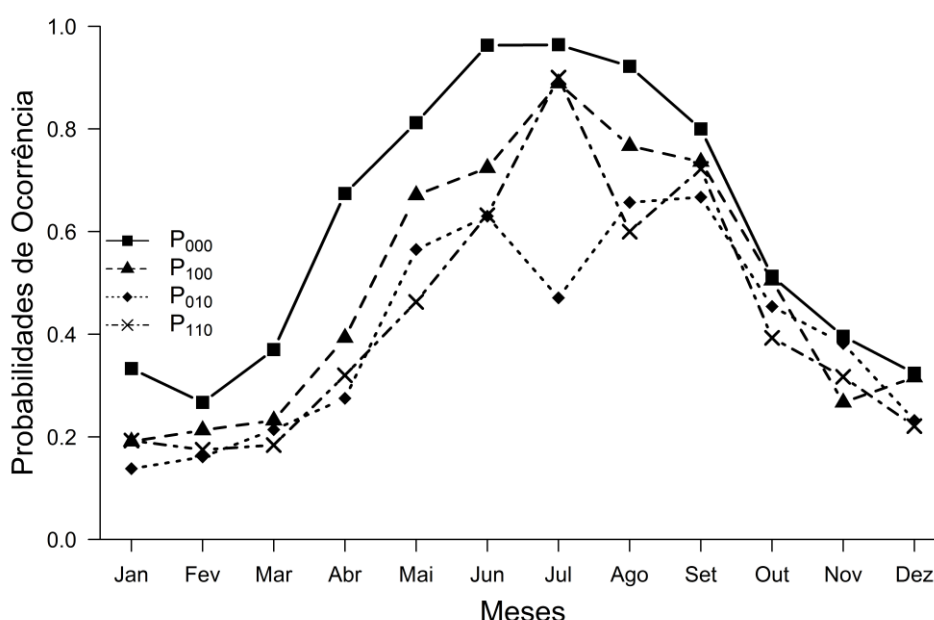


Figura 4. Probabilidades de transição estimadas pela cadeia de Markov de 2ª ordem para a não ocorrência de precipitação no município de Ji-Paraná: P_{000} (dia seco precedido de dois dias secos), P_{100} (dia seco precedido por um dia seco e outro chuvoso), P_{010} (dia seco precedido por um dia chuvoso e outro seco) e P_{110} (dia seco precedido de dois dias chuvosos).

FONTE: elaborado em ambiente R

Outro fator merece ser analisado, nas probabilidades de dia seco antecedido por dia chuvoso e dia seco respectivamente (P_{010}), pois em julho houve uma descontinuidade no crescimento desta probabilidade. Descritivamente pode-se considerar que um dia seco neste mês tem menores chances de acontecer se precedido por um dia chuvoso anterior e outro seco.

Houve uma quebra no sentido das probabilidades no mês de agosto, para dia seco precedido por dois dias chuvosos, que diminuiu drasticamente em relação ao mês anterior, que havia fornecido maior chance de não ocorrer precipitação se antecedido por

dois dias de chuva. Este comportamento se repete quando se trata da probabilidade do dia atual ser chuvoso (Figura 5) só que de maneira análoga.

Portanto, quando o dia atual é precedido de dois dias secos consecutivos, suas probabilidades serão menores quando comparadas com as demais probabilidades de transição para dia chuvoso.

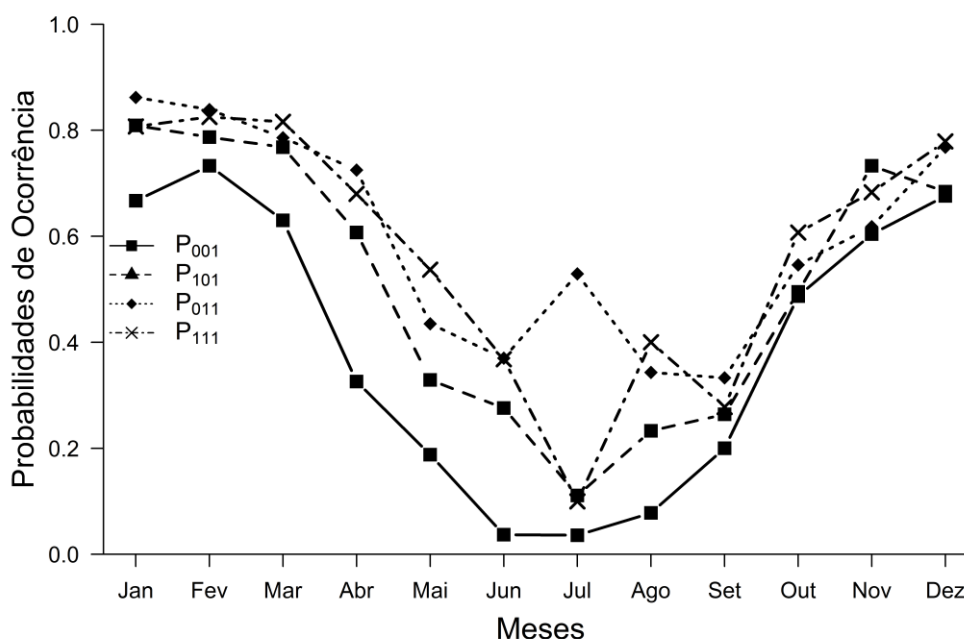


Figura 5. Probabilidades de transição estimadas pela cadeia de Markov de 2ª ordem para a não ocorrência de precipitação no município de Ji-Paraná: P_{001} (dia chuvoso precedido de dois dias secos), P_{101} (dia chuvoso precedido por um dia seco e outro chuvoso), P_{111} (dia chuvoso precedido por um dia chuvoso e outro seco) e P_{011} (dia chuvoso precedido de dois dias chuvosos)

FONTE: elaborado em ambiente R

Em janeiro houve maior probabilidade de acontecer um dia chuvoso quando o dia passado foi chuvoso e o imediatamente anterior a este foi seco, ou seja, é bem provável que se houver um dia sem e outro com chuva o próximo dia terá 0,862 de probabilidade de ocorrer precipitação acima de 1mm em janeiro.

As probabilidades de transição P_{011} , P_{101} e P_{111} foram próximas em quase todos os meses do ano, onde apenas entre os meses de seca no Município houve maiores diferenças entre estas probabilidades, principalmente no mês de julho onde foi relativamente maior em relação as demais, o que indica uma maior chance de acontecer um dia chuvoso se o anterior também foi chuvoso e outro seco.

Definidas as probabilidades de transição das ordens 1 e 2 da cadeia de Markov para ocorrência de dias secos e chuvosos, foi realizada uma investigação das melhores

distribuições de probabilidades que ajustam as quantidades de precipitação pluvial do Município de Ji-Paraná (Tabela 5).

Tabela 5. Parâmetros de ajuste das distribuições de probabilidades Exponencial, Gama e Weibull para as médias de precipitação diárias de cada mês e teste de Kolmogorov-Smirnov para aderência das distribuições estatísticas aos dados.

distribuições estatísticas aos dados.

Meses	Parâmetros					Teste de Kolmogorov-Smirnov					
	Exponencial	Gama		Weibull		Exponencial		Gama		Weibull	
	λ	γ	β	α	β	D	p-valor	D	p-valor	D	p-valor
Jan.	0,092	14,082	1,297	3,531	12,035	0,521	0,000	0,228	0,362	0,210	0,458
Fev.	0,074	24,576	1,823	5,155	14,627	0,507	0,000	0,107	0,988	0,125	0,950
Mar.	0,097	21,218	2,057	5,060	11,230	0,500	0,001	0,167	0,735	0,170	0,721
Abr.	0,147	8,025	1,176	3,370	7,607	0,371	0,023	0,168	0,734	0,127	0,943
Mai.	0,437	2,838	1,240	1,726	2,587	0,298	0,111	0,191	0,578	0,193	0,564
Jun.	1,11	*	*	*	*	0,333	0,071	*	*	*	*
Jul.	2,319	*	*	*	*	0,302	0,129	*	*	*	*
Ago.	0,956	0,703	0,672	0,796	0,926	0,204	0,500	0,129	0,937	0,138	0,900
Set.	0,495	2,018	0,999	1,559	2,241	0,257	0,232	0,206	0,487	0,226	0,369
Out.	0,18	7,380	1,331	3,053	6,204	0,377	0,020	0,113	0,978	0,150	0,838
Nov.	0,148	10,217	1,509	4,159	7,485	0,377	0,020	0,212	0,446	0,165	0,747
Dez.	0,108	18,317	1,978	5,305	10,087	0,470	0,001	0,163	0,762	0,128	0,939

* As distribuições de probabilidades testadas não tiveram ajuste

FONTE: Elaboração através do *software* R

Foram testadas as distribuições estatísticas Exponencial, Gama e Weibull, e pelo teste de Kolmogorov-Smirnov a 5% de significância verificou-se a aderência do ajuste. As distribuições Gama e Weibull se ajustaram a todos os meses, exceto junho e julho. Isto ocorre porque nelas há o pressuposto de todas as observações serem maiores que zero. Como alternativa, a distribuição Exponencial se ajustou a estes meses, assim como pelo teste de KS ajustou-se para maio, agosto e setembro.

Todas as distribuições testadas se ajustaram para maio, agosto e setembro, mas a distribuição Exponencial foi a que mais superestimou os valores reais, que pode ser observado pela estatística D de Kolmogorov-Smirnov. As distribuições Gama e Weibull se aderiram ao período de agosto a maio, onde contém a época de chuvas do ano, e somente a Exponencial aderiu a precipitação dos meses de junho e julho.

Portanto, são apresentadas na Tabela 6 as distribuições com melhor aderência pelo teste de Kolmogorov-Smirnov a 5% de significância e seus parâmetros e a ordem ótima obtida pelos critérios de AIC e BIC.

Tabela 6. Resumo das distribuições de probabilidades com melhor aderência e ordem mais apropriada para a ocorrência de dias secos e chuvosos.

Meses	Distribuição de Probabilidades	Parâmetros		Critérios	
		Parâmetro 1	Parâmetro 2	AIC	BIC
Jan.	Weibull	3,531 ⁽⁴⁾	12,035 ⁽⁵⁾	2	2
Fev.	Gama	24,576 ⁽²⁾	1,823 ⁽³⁾	2	2
Mar.	Gama	21,218 ⁽²⁾	2,057 ⁽³⁾	2	2
Abr.	Weibull	3,37 ⁽⁴⁾	7,607 ⁽⁵⁾	2	2
Mai.	Gama	2,838 ⁽²⁾	1,24 ⁽³⁾	2	2
Jun.	Exponencial	1,11 ⁽¹⁾	-	2	2
Jul.	Exponencial	2,319 ⁽¹⁾	-	2	2
Ago.	Gama	0,703 ⁽²⁾	0,672 ⁽³⁾	2	2
Set.	Gama	2,018 ⁽²⁾	0,999 ⁽³⁾	2	2
Out.	Gama	7,38 ⁽²⁾	1,331 ⁽³⁾	2	2
Nov.	Weibull	4,159 ⁽⁴⁾	7,485 ⁽⁵⁾	2	2
Dez.	Weibull	5,305 ⁽⁴⁾	10,087 ⁽⁵⁾	2	2

1 = parâmetro de forma λ , 2 = parâmetro de forma γ , 3 = parâmetro de escala β ,

4 = parâmetro de posição α , 5 = parâmetro de controle da curva β

FONTE: Elaboração através do *software* R

Com isso, verificou-se pela estatística D do teste de KS (Tabela 5), a qual representa os maiores valores das diferenças entre os dados originais e os ajustados, e assim foi escolhida a distribuição com base no menor valor desta diferença. Sendo assim a distribuição Gama ajustou-se melhor para os meses de fevereiro, março, maio, agosto, setembro e outubro para as estimativas das quantidades de precipitação de Ji-paraná. A distribuição Weibull é mais adequada para os meses de janeiro, abril, novembro e dezembro. Já a distribuição Exponencial adequa-se mais aos meses de seca do Município, junho e julho.

No âmbito deste trabalho verificou-se também uma unanimidade em relação a segunda ordem da cadeia de Markov para transição entre dias secos e chuvosos, sendo que esta ordem apresentou menores valores do critério de AIC e também do BIC.

Sendo assim, as cadeias de Markov exercem grande força descritiva para compreensão da ocorrência de chuvas, mas além disto, como relata Keller Filho *et al.* (2006), as cadeias de Markov podem ser utilizadas para simulação de sequências de dias secos e chuvosos, que podem ser empregadas a situações com poucas observações.

5. CONCLUSÕES E CONSIDERAÇÕES

Para fins descritivos as probabilidades de transição de primeira e segunda ordem da cadeia de Markov de dois estados tiveram bom comportamento para caracterização das quantidades de dias secos e chuvosos, mas pelos critérios estabelecidos neste trabalho a segunda ordem da cadeia de Markov foi considerada ótima para a ocorrência de chuvas no Município de Ji-Paraná, ou seja, dois dias são classificados como adequados para se prever as chances de ocorrer dias chuvosos ou secos na região.

Para determinação do volume de precipitação a distribuição probabilidades Gama é mais adequada para os meses de fevereiro, março, maio, agosto, setembro e outubro, a Weibull para os meses de janeiro, abril, novembro e dezembro e a Exponencial para os meses junho e julho. Não foi possível calcular os parâmetros da Gama e Weibull nos meses de junho e julho, mas nos demais seus parâmetros foram significativos. A distribuição Exponencial foi ajustada a todos meses, mas seus parâmetros não foram significantes na maioria, exceto nos meses de estiagem no Estado de Rondônia.

Com base nas cadeias de Markov e em conjunto com as distribuições estatísticas de probabilidades usadas, poderão ser realizados novos trabalhos para simular de séries de dados com poucas observações no Município de Ji-Paraná.

REFERÊNCIAS

- AGRITEMPO - Sistema de Monitoramento Agrometeorológico. **Estações meteorológicas para o estado de Rondônia.** Disponível em: <<https://www.agritempo.gov.br/agritempo/jsp/Estacao/index.jsp?siglaUF=RO>>. Acesso em: 23 de março de 2017.
- AKAIKE, H., 1974: A new look at the statistical model identification. **IEEE Trans. Autom. Control**, 19, 716–723.
- ALMEIDA, R. M. B. **Características climatológicas do regime de chuva em Minas Gerais.** Tese de Mestrado, Viçosa, UFV, 1995, 64p.
- ANDRADE JÚNIOR, A. S.; FRIZZONE, J. A.; SENTELHAS, P. C. Simulação da precipitação diária para Parnaíba e Teresina, PI, em planilha eletrônica. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande, PB, DEAg/UFPB, v.5, n.2, p.271-278, 2001. Disponível em: <<http://www.agriambi.com.br>>.
- ANDRADE, A. R. S.; FREITAS, J. C.; BRITO, J. I. B.; GUERRA, H. O. C.; XAVIER, J. F. Aplicação da probabilidade condicional e do processo de cadeia de Markov na análise da ocorrência de períodos secos e chuvosos para o município de Garanhuns, PE, Brasil. **Ambi-Agua**, Taubaté, v. 4, n. 1, p. 169-182, 2009.
- ASSIS, F. N.; ARRUDA, H. V.; PEREIRA, A. R. **Aplicações de estatística à climatologia: teoria e prática.** Pelotas: Ed. Universitária/UFPel, 161p. ,1996.
- ASSIS, F. N.; VILLA NOVA, N. A. Modelagem da ocorrência e da quantidade de chuva em Piracicaba. **Sci. agric.**, Piracicaba, 172-182p., set./dez. 1994
- AZEVEDO, P. V.; LEITÃO, M. M. V. B. R. Aplicação de processos em cadeia de Markov as precipitações diárias no estado da Paraíba. **Revista Brasileira de Meteorologia**: 1990; Vol. 5(1), 389-402.
- BACK, A. J.; GALATTO, S. L.; COAN, B. P.; DELLA, J. P. Probabilidade de ocorrência de períodos secos e chuvosos em Saudades-SC. **XII Simpósio Ítalo-Brasileiro de Engenharia Sanitária e Ambiental – SIBESA**, Associação Brasileira de Engenharia Sanitária e Ambiental (ABES), Natal, RN, 19 a 21 de maio de 2014.
- BARKOTULLA, M. A. B. Stochastic Generation of the Occurrence and Amount of Daily Rainfall. **Pak.j. Stat. Oper. Res.** v.6 n.1, p. 61-73, Bangladesh, 2010.
- BASAK, P. On the Markov chain models for monssonal rainfall occurrence in different zones of West Bengal. **India Journal of Radio & Space Physics**, v. 43, p. 349-254, December 2014.
- BAÚ, A. L.; AZEVEDO, C. A. V.; BRESOLIN, A. A. Modelagem da precipitação pluvial diária intra-anual da Bacia Hidrográfica Paraná III associada aos eventos ENOS. **R. Bras. Eng. Agric. Ambiental**, v.17, n.8, p.883–891, 2013.
- BEZERRA, U. A.; FARIAS, C. A. S.; SILVA FILHO, J. A. Cadeias de Markov para estimativa diária de ocorrência de vazões no rio Piancó. **Congresso Técnico Científico**

da Engenharia e da Agronomia – CONTECC’ 2015, 15 a 18 de setembro de 2015 - Fortaleza-CE, Brasil.

CAMPOS, H. **Estatística experimental não-paramétrica**. 3ª ed., ESALQ. Piracicaba - SP, 1979.

CATALUNHA, M. J. **Avaliação do ajustamento de funções densidade de probabilidade a séries de precipitação pluvial no Estado de Minas Gerais**. 72f. Dissertação (Mestrado em Meteorologia Agrícola) - Universidade Federal de Viçosa, Viçosa. 2000.

CATALUNHA, M. J.; SEDIYAMA, G. C.; LEAL, B. G.; SOARES, C. P. B.; RIBEIRO, A. Aplicação de cinco funções densidade de probabilidade a séries de precipitação pluvial no Estado de Minas Gerais. **Revista Brasileira de Agrometeorologia**, Santa Maria, v. 10, n. 1, p. 153-162, 2002.

COE, R.; STERN, R.D. Fitting models to daily rainfall data. **Journal of Applied Meteorology**, 21:1024-31, 1982.

CRUCIANI, D.E.; MACHADO, R.E.; SENTELHAS, P. C. Modelos da distribuição temporal de chuvas intensas em Piracicaba, SP. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 6, n. 1, p. 76-82, 2002.

DELIGNETTE-MULLER, M. L.; DUTANG, C. **fitdistrplus: An R Package for Fitting Distributions**. Journal of Statistical Software, 64(4), 1-34, 2015. URL: <<http://www.jstatsoft.org/v64/i04/>>.

DETZEL, D. H. M. **Geração de séries sintéticas de precipitação**. Dissertação de Mestrado, Departamento de Engenharia Hidráulica e Saneamento, Universidade Federal do Paraná, Curitiba, p 197, 2009.

FRANCA, R. R. Climatologia das chuvas em Rondônia—período 1981-2011. **Revista Geografias**, v. 1, n. 20, p. 44-58, 2015.

FURLAN, D. N. **Variabilidade temporal e espacial das chuvas e do balanço hídrico no estado de Rondônia: caracterização e análise de tendência**. Dissertação (Mestrado) - Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, 2009.

GORDON, P. **Théorie des chaines de Markov finies et ses applications**. Paris: Dunod, 130p. 1965.

KELLER FILHO, T.; ZULLO JUNIOR, J.; LIMA, P. R. S. R. Análise da transição entre dias secos e chuvosos por meio da cadeia de Markov de terceira ordem. **Pesquisa Agropecuária Brasileira**, Brasília, v.41, n.9, p.1341-1349, set. 2006.

MACHADO, M. B.; DAMÉ, R. C. F.; TEIXEIRAGANDRA, C. F. A.; GUEDES, H. A. S.; COUTO, R. S. Modelos probabilísticos ajustados a dados de precipitação máxima diária anual observada e gerada. **XLIII Congresso Brasileiro de Engenharia Agrícola - CONBEA 2014**, Centro de Convenções “Arquiteto Rubens Gil de Camillo” - Campo Grande –MS, 27 a 31 de julho de 2014.

MOURA, I. J. M.; SANTOS, D. F.; PINHEIRO, F. G. M.; OLIVEIRA, C. J. Caracterização dos Períodos Seco e Chuvoso da Cidade de Fortaleza (CE). **Revista do Centro de Ciências Naturais e Exatas - UFSM**. Ciência e Natura, Santa Maria, v. 37 Ed. Especial SIC, 2015, p. 03-07.

NETO, D. D.; ASSIS, J.; TIMM, L. C.; MANFRON, P. A.; SPAROVEK, G.; MARTIN, T. N. Ajuste de Modelos de Distribuição de Probabilidade a Séries Históricas de Precipitação Pluvial Diária em Piracicaba- SP. **Rev. Bras. Agrometeorologia**, v. 13, n. 2, p. 273-283, 2005.

NISHIJIMA, T. **Modelagem Markoviana da precipitação pluvial diária e simulação do rendimento esperado de soja no município de Cruz Alta (RS)**. Tese (Doutorado) Submetida ao Programa de Pós-graduação em Recursos Hídricos e Saneamento Ambiental da Universidade Federal do Rio Grande do Sul, UFRS, Porto Alegre, RS, dezembro de 2004.

PAIVA, E. M. C. D.; CLARKE, R. T. Modelagem estocástica da precipitação na Amazônia. **Revista Brasileira de Recursos Hídricos - RBRH**, v. 2, n. 2, p. 157-171, Jul/Dez 1997.

PAULA JÚNIOR, G. F. **Ajustamento de modelos de probabilidade para variáveis de precipitação na cidade de Cacoal/RO**. Trabalho de Conclusão de Curso de Estatística - Universidade Federal de Rondônia, Ji-Paraná, 29 p., 2016.

PUNYAWARDENA, B.V.R.; KULASIRI, D. **On development and comparative study of two Markov models of rainfall in the Dry Zone of Sri Lanka**. Centre for Computing and Biometrics, Lincoln Universit, Research Report No:96/11, October 1996.

R CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2017. URL: <<https://www.R-project.org/>>.

ROCHA FILHO, J. V.; PINTO, E. J. A. Geração estocástica de séries de precipitação com cadeias de Markov híbrida de segunda ordem. **XXI Simpósio Brasileiro de Recursos Hídricos**, Brasília-DF, 22 a 27 de novembro de 2015.

RODRIGUES, J. A.; FILHO, J. S.; CHAVES, L. M. Funções densidade de probabilidade para a estimativa de precipitação mensal. **Semina: Ciências Exatas e Tecnológicas**, Londrina, v. 34, n. 1, p. 03-08, jan. /jul. 2013.

SCHOOF, J. T.; PRYOR, S. C. On the Proper Order of Markov Chain Model for Daily Precipitation Occurrence in the Contiguous United States. **Journal of Applied Meteorology and Climatology**, American Meteorological Society, v. 47, p. 2477-2486, September 2008.

SCHWARZ, G., 1978: Estimating the dimension of a model. **Ann. Stat.** 6, 461–464.

SEDIYAMA, G. C.; MELO, J. S. P.; ALVES, A. R.; COELHO, D. T. Dependência temporal e espacial das probabilidades de dias secos e, ou, chuvosos em dias consecutivos, para intervalos mensais. **Revista Ceres**, 43(247):267-280, 1996.

SHIAU, J. T; MORRADES, R. Copula-Based drought severity-duration-frequency analysis in Iran. **Meteorological Applications**, Malden, v. 16, n.4, p. 481-489, 2009.

VIANA, T. V. A.; AZEVEDO, B. M.; BOMFIM, G. V.; ANDRADE JÚNIOR, A. S. Probabilidade de ocorrência de períodos secos e chuvosos em Pentecoste, CE. **Irriga**, Botucatu, v. 7, n. 3, p. 226-229, 2002.

WILKS, D.S. **Statistical methods in the atmospheric sciences**. International Geophysics Series, ELSEVIER, Amsterdam, 2 ed., 2006.

ZANETTI, S. S.; OLIVEIRA, V. P. S.; PRUSKI, F. F. Validação do Modelo ClimaBR em Relação ao Número de Dias Chuvosos e à Precipitação Total Diária. **Eng. Agric.**, Jaboticabal, v.26, n.1, p.96-102, jan./abr. 2006.

ZUCCHINI W.; ADAMSON P. T. The occurrence and severity of droughts in South Africa, **Water Research Commission**, Pretoria, 91/1/94, 1984.

APÊNDICE A – ROTINA DE ANÁLISE EM AMBIENTE R

```
#### OBTENÇÃO DAS ESTIMATIVAS DA CADEIA DE MARKOV ###
```

```
## Diretório de trabalho para exportar dados
```

```
setwd("~/Trabalho_de_Conclusao_de_Curso/Analise/JIPA")
```

```
#### data.frame das frequencias de ocorrência de chuva
```

```
frequencias = read.csv2(file =
```

```
"~/Trabalho_de_Conclusao_de_Curso/Analise/JIPA/DADOS/Nij-Nijk.csv", header = T,  
dec = ",")
```

```
### Função responsável por criar a Cadeia de markov de 1ª ordem
```

```
Markov.1 <- function( x ){
```

```
  ## Esta função receberá as frequencia de ocorrencia, sendo que:
```

```
  ##                linha terá os valores das frequencias mensais
```

```
  ##                coluna terá o tipo de frequencia conforme as cadeias
```

```
  m = 1 # ordem da cadeia
```

```
  s = 2 # estados
```

```
  colunas = cbind()
```

```
  c = r = 1
```

```
  while (r <= dim( x )[1] ){
```

```
    N00 = x[r, "N00"]; N01 = x[r, "N01"]
```

```
    N10 = x[r, "N10"]; N11 = x[r, "N11"]
```

```
    N0 = N00 + N01; N1 = N10 + N11
```

```
    p00 = N00/N0
```

```
    p01 = N01/N0
```

```
    p10 = N10/N1
```

```
    p11 = N11/N1
```

```
    L <- N00*log( p00 ) + N01*log( p01 ) + N10*log( p10 ) + N11*log( p11 )
```

```
    AIC <- -2*L + 2*(s^( m ))*(s-1)
```

```
    BIC <- -2*L + (s^( m ))*(log( N0+N1 ))
```

```
    linhas <- cbind(p00, p10, p01, p11, AIC, BIC)
```

```
    colunas <- rbind(colunas, linhas)
```

```
    r = r + 1
```

```
  }
```

```
  return( colunas )
```

```
}
```

```
### Função responsável por criar a Cadeia de markov de 2ª ordem
```

```
Markov.2 <- function( x ){
```

```
  ## Esta função receberá as frequencia de ocorrencia, sendo que:
```

```
  ##                linha terá os valores das frequencias mensais
```

```
  ##                coluna terá o tipo de frequencia conforme as cadeias
```

```
  m = 2 # ordem da cadeia
```

```
  s = 2 # estados
```

```
  colunas = cbind( )
```

```
  c = r = 1
```

```
  while (r <= dim(x)[1]){
```

```

N000 = x[r, "N000"]; N010 = x[r, "N010"]; N011 = x[r, "N011"]
N001 = x[r, "N001"]; N100 = x[r, "N100"]; N101 = x[r, "N101"]
N110 = x[r, "N110"]; N111 = x[r, "N111"]

N00 = N000+N001; N01 = N010+N011; N10 = N100+N101; N11 = N110+N111
p000 = N000/N00; p001 = N001/N00
p010 = N010/N01; p011 = N011/N01
p100 = N100/N10; p101 = N101/N10
p110 = N110/N11; p111 = N111/N11
L <- N000*log( p000 ) + N010*log( p010 )+ N001*log( p001 ) + N110*log( p110 ) +
N100*log( p100 ) + N111*log( p111 )
AIC <- -2*L + 2*(s^( m ))*(s-1)
BIC <- -2*L + (s^( m ))*(log( N00+N01+N10+N11 ))
linhas <- cbind(p000, p100, p010, p110, p001, p101, p011, p111, AIC, BIC)
colunas <- rbind(colunas, linhas)
r =r+1
}
return( colunas )
}

```

```
#### Gerando a cadeia de markov
```

```

pij = round( as.data.frame( Markov.1( frequencias[, 2:13])), digits = 3 )
pijk = round( as.data.frame( Markov.2( frequencias[, 2:13])), digits = 3 )

```

```
### nomeando os data.frames para exportar
```

```

rownames( pij ) <- rownames( pijk ) <- c("Jan", "Fev", "Mar", "Abr", "Mai", "Jun",
"Jul", "Ago", "Set", "Out", "Nov", "Dez")

```

```
## exporta para a pasta do diretorio
```

```

write.table(pij, file = "Pij.csv", dec=".", sep = ";",
quote = FALSE, col.names=T,row.names = F) #### Matriz de 1ª ordem
write.table(pijk, file = "Pijk.csv", dec=".", sep = ";",
quote = FALSE, col.names=T,row.names = F) #### Matriz de 2ª ordem
#### FIM DA GERAÇÃO DAS CADEIAS DE MARKOV ####

```

```
#-----#
```

```
#### AJUSTAMENTO DE DISTRIBUIÇÃO DE PROBABILIDADES ####
```

```
## Diretório de trabalho para exportar dados
```

```
setwd( "~/_Trabalho_de_Conclusao_de_Curso/Analise/JIPA/DADOS" )
```

```
#### Pacotes de funções do R
```

```

library( ggplot2 )
library( fitdistrplus )
library( logspline )

```

```
### Função para ajuste de distribuição de probabilidades
```

```

ajuste <- function(meses, dist, botao, dados_mensais){
## São necessários que tenha uma lista com os nomes dos meses

```

```

## Que escreva o nome da distribuição: "exp", "gamma", "weibull"
## que indique a opção: T para quando usar a gamma e weibull e
##          F para quando usar a exponencial
## e precisa do data.frame com os dados mensais no tipo: coluna com os meses
##          linha com os valores dos anos
parametros <- tabela<- c( )
p <- paste("p", dist, sep = "")
for ( i in meses ){
  prp <- as.numeric( dados_mensais[ , i ] )
  cond <- TRUE
  if ( botao ){ for ( j in prp ){
    if ( j == 0 ){ cond <- FALSE }
  }}
  if ( cond ) {
    Fdist = fitdist( prp, dist )
    parametros <- cbind( Fdist$estimate )
    if ( dist != "exp" ){
      D <- as.numeric( ks.test( prp, p, Fdist$estimate[1], Fdist$estimate[2] )$statistic )
      p.valor <- ks.test( prp, p, Fdist$estimate[1], Fdist$estimate[2] )$p.value
    } else {
      D <- as.numeric( ks.test( prp, p, Fdist$estimate[1] )$statistic )
      p.valor <- ks.test( prp, p, Fdist$estimate[1] )$p.value
    }
    AIC <- Fdist$aic
    BIC <- Fdist$bic
    tabela = rbind( tabela, c( parametros, D, p.valor, AIC, BIC ) )
    colnames(tabela) <- c( rownames(parametros), "D", "p-valor", "AIC", "BIC")
  } else { parametros[1] = parametros[2] = D = p.valor = AIC = BIC = NA
    tabela = rbind( tabela, c(parametros, D, p.valor, AIC, BIC))
    colnames(tabela) <- c( "shape", "rate", "D", "p-valor", "AIC", "BIC") }
}
rownames( tabela ) <- meses
return( round(( tabela ), 3) )
}

```

```
## lista com meses
```

```
meses <- c( "Jan", "Fev", "Mar", "Abr", "Mai", "Jun",
            "Jul", "Ago", "Set", "Out", "Nov", "Dez" )
```

```
## data.frame com os dados de precipitação mensal
```

```
AJ <- read.csv2( file =
  "~/_Trabalho_de_Conclusao_de_Curso/Analise/JIPA/DADOS/Media_periodo.csv",
  header = T, dec = "," )
```

```
## Valores dos ajustes das distribuições
```

```
ajuste(meses, "gamma", T, AJ[,-1]) ## GAMA
```

```
ajuste(meses, "weibull", T, AJ[,-1]) ## WEIBULL
```

```
ajuste(meses, "exp", F, AJ[,-1]) ## EXPONENCIAL
```

```
#### FIM DO AJUSTE ####
```

```
#-----#
```

```
#### GRÁFICOS UTILIZADOS ####
```

```
#### Pacotes de funções do R
```

```
require(ggplot2)
```

```
## Diretório de trabalho para exportar dados
```

```
setwd( "~/_Trabalho_de_Conclusao_de_Curso/Analise/JIPA" )
```

```
#### Dados utilizados
```

```
pij <- read.csv2(file = "~/_Trabalho_de_Conclusao_de_Curso/Analise/JIPA/Pij.csv",  
header=T, dec = ",") #### transição de 1ª ordem
```

```
pijk <-read.csv2(file = "~/_Trabalho_de_Conclusao_de_Curso/Analise/JIPA/Pijk.csv",  
header = T, dec = ",") #### transição de 2ª ordem
```

```
qtd <- read.csv2(file =  
"~/_Trabalho_de_Conclusao_de_Curso/Analise/JIPA/DADOS/Seco-Chuvoso.csv",  
header = T, dec = ",") #### dias secos e chuvosos
```

```
## Número de dias seco e chuvosos da serie
```

```
M <- meses <- c("Jan", "Fev", "Mar", "Abr", "Mai", "Jun",  
"Jul", "Ago", "Set", "Out", "Nov", "Dez",  
"Jan", "Fev", "Mar", "Abr", "Mai", "Jun",  
"Jul", "Ago", "Set", "Out", "Nov", "Dez")
```

```
Mes<-seq(1, 12, 1)
```

```
Me = as.numeric( rbind( c(Mes, Mes) ))
```

```
Tempo <- as.character(rbind(c( rep("Seco", 12), rep("Chuvoso", 12))))
```

```
Precipitacao <-as.numeric(rbind(c(as.numeric(qtd$Seco), as.numeric(qtd$Chuvoso))))
```

```
dat <- data.frame( Tempo, Precipitacao, M, Me)
```

```
#### Graficos das quantidades de dias secos e chuvosos do periodo
```

```
grafico <- ggplot(dat, aes(x = as.factor(Me), y = as.numeric(Precipitacao),  
fill=as.factor(Tempo)))
```

```
x11(); grafico + geom_bar(stat="identity", position = "dodge",  
aes(fill = Tempo, group = Tempo))+
```

```
scale_fill_manual(values = c( "blue", "firebrick2"), name = "Tempo") +
```

```
xlab("Meses") + # adiciona descrição do eixo x
```

```
ylab("Número de dias") + # adiciona descrição do eixo y
```

```
ggtitle("")+
```

```
scale_x_discrete(breaks=as.factor(Me), labels = M)
```

```
#### conjunto de funções para fazer o grafico de dependencia temporal
```

```
x= c(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12)
```

```

P00 <- pij$p00; P10 <- pij$p10; P01 <- pij$p01; P11 <- pij$p11
P000 <- pijk$p000; P001 <- pijk$p001; P010 <- pijk$p010
P011 <- pijk$p011; P100 <- pijk$p100; P101 <- pijk$p101
P110 <- pijk$p110; P111 <- pijk$p111
locais=c(expression(P[0][0]), expression(P[1][0]),
           expression(P[0][1]), expression(P[1][1]))
locais1=c(expression(P[0][0][0]), expression(P[1][0][0]),
           expression(P[0][1][0]), expression(P[1][1][0]))
locais2=c(expression(P[0][0][1]), expression(P[1][0][1]),
           expression(P[0][1][1]), expression(P[1][1][1]))

#### Exportar grafico de dependencia temporal de 1 dia
tiff(filename = "Grafico_de_Dependencia_Temporal_de_1_dia.png",
      width = 18, height = 12, units = "cm",
      compression = c("lzw+p"),
      bg = "white", res = 1000, family = "", restoreConsole = TRUE,
      type = c("windows"))
par(family = "arial")
m2 = matrix(c(1,2,1,1), nrow = 2, ncol = 2, byrow = TRUE)
layout(m2)
layout.show(2)
par(mar = c(5,4,2,.5), mgp = c(2.7,1,0), mai = c(0.9,.9,0.2,0.2))
## Esboço de Referencia
plot(P00~x, xaxt = "n", ylim = c(0,1), ylab = "Probabilidades de Ocorrência",
     yaxs = "i", type = "b",
     xlab = "Meses", las = 2, bty = "l", col = "black", pch = 19, lwd = 2,
     cex.axis = 1.1, cex.lab = 1.5, cex = 1.5)
axis(side = 1, at = 1:12, cex.axis = 1.1, labels = meses, line = 0, las = 1)
par(new = TRUE)## inserção de gráfico
plot(P10~x, yaxt = "n", xaxt = "n", ylim = c(0,1), ylab = "", yaxs = "i",
     type = "b", col = "red",
     pch = 19, xlab = "", las = 2, bty = "l", lwd = 2, cex.axis = 1.1,
     cex.lab = 1.3, cex = 1.5)
leg.txt = c("P00", "P10", "P10", "P11")
legend("left", legend = locais, pch = c(19,19,NA,NA), cex = 1.2,
      bty = "n", lty = c(1,1,NA,NA), bg = "n", col = c("black","red",NA,NA))
par(new = TRUE)## inserção de gráfico
plot(P01~x, yaxt = "n", xaxt = "n", ylim = c(0,1), ylab = "", yaxs = "i",
     type = "b", col = "blue",
     pch = 19, xlab = "", las = 2, bty = "l", lwd = 2, cex.axis = 1.1,
     cex.lab = 1.3, cex = 1.5)
legend("left", legend = locais, pch = c(19,19,19,NA), cex = 1.2,
      bty = "n", lty = c(1,1,1,NA),bg = "n", col = c("black","red","blue",NA))
par(new = TRUE)## inserção de gráfico
plot(P11~x, yaxt = "n", xaxt = "n", ylim = c(0,1), ylab = "", yaxs = "i",

```

```

    type = "b", col = "green",
    pch = 19, xlab = "", las = 2, bty = "l", lwd = 2, cex.axis = 1.1,
    cex.lab = 1.3, cex = 1.5)
legend("left", legend = locais, pch = c(19,19,19,19), cex = 1.2,
      bty = "n", lty = c(1,1,1,1), bg = "n", col = c("black","red","blue","green"))
dev.off()

#### Exportar grafico de dependencia temporal de 2 dias (SECO)
tiff(filename = "Grafico_de_Dependencia_Temporal_de_2_dia(SECO).png",
      width = 18, height = 12, units = "cm",
      compression = c("lzw+p"),
      bg = "white", res = 1000, family = "", restoreConsole = TRUE,
      type = c("windows"))
par(family="arial")
m2 = matrix(c(1,2,1,1), nrow = 2, ncol = 2, byrow = TRUE)
layout(m2)
layout.show(2)
par(mar = c(5,4,2,.5), mgp = c(2.7,1,0), mai = c(0.9,.9,0.2,0.2))
## Esboço de Referencia
plot(P000~x, xaxt = "n", ylim = c(0,1), ylab = "Probabilidades de Ocorrência",
     yaxs = "i", type = "b",
     xlab = "Meses", las = 2, bty = "l", col = "black", pch = 19, lwd = 2,
     cex.axis = 1.1, cex.lab = 1.5, cex = 1.5)
axis(side = 1, at = 1:12, cex.axis = 1.1, labels = meses, line = 0, las = 1)
par(new = TRUE)## inserção de gráfico
plot(P100~x, yaxt = "n", xaxt = "n", ylim = c(0,1), ylab = "", yaxs = "i",
     type = "b", col = "red",
     pch = 19, xlab = "", las = 2, bty = "l", lwd = 2, cex.axis = 1.1,
     cex.lab = 1.3, cex = 1.5)
leg.txt = c("P00","P10","P10","P11")
legend("left", legend = locais1, pch = c(19,19,NA,NA), cex = 1.2,
      bty = "n", lty = c(1,1,NA,NA), bg = "n", col = c("black","red",NA,NA))
par(new = TRUE)## inserção de gráfico
plot(P010~x, yaxt = "n", xaxt = "n", ylim = c(0,1), ylab = "", yaxs = "i",
     type = "b", col = "blue",
     pch = 19, xlab = "", las = 2, bty = "l", lwd = 2, cex.axis = 1.1,
     cex.lab = 1.3, cex = 1.5)
legend("left", legend = locais1, pch = c(19,19,19,NA), cex = 1.2,
      bty = "n", lty = c(1,1,1,NA), bg = "n", col = c("black","red","blue",NA))
par(new = TRUE)## inserção de gráfico
plot(P110~x, yaxt = "n", xaxt = "n", ylim = c(0,1), ylab = "", yaxs = "i",
     type = "b", col = "green",
     pch = 19, xlab = "", las = 2, bty = "l", lwd = 2, cex.axis = 1.1,
     cex.lab = 1.3, cex = 1.5)
legend("left", legend = locais1, pch = c(19,19,19,19), cex = 1.2,

```

```

    bty = "n", lty = c(1,1,1,1), bg = "n", col = c("black","red","blue","green"))
dev.off()

#### Exportar grafico de dependencia temporal de 2 dias (CHUVA)
tiff(filename = "Grafico_de_Dependencia_Temporal_de_2_dia(CHUVA).png",
    width = 18, height = 12, units = "cm",
    compression = c("lzw+p"),
    bg = "white", res = 1000, family = "", restoreConsole = TRUE,
    type = c("windows"))
par(family = "arial")
m2 = matrix(c(1, 2, 1, 1), nrow = 2, ncol = 2, byrow = TRUE)
layout(m2)
layout.show(2)
par(mar = c(5, 4, 2, .5), mgp = c(2.7, 1, 0), mai = c(0.9, .9, 0.2, 0.2))
## Esboço de Referencia
plot(P001~x, xaxt = "n", ylim = c(0,1), ylab = "Probabilidades de Ocorrência",
    yaxs = "i", type = "b",
    xlab = "Meses", las = 2, bty = "l", col = "black", pch = 19, lwd = 2,
    cex.axis = 1.1, cex.lab = 1.5, cex = 1.5)
axis(side = 1, at = 1:12, cex.axis = 1.1, labels = meses, line = 0, las = 1)
par(new = TRUE) ## inserção de gráfico
plot(P101~x, yaxt = "n", xaxt = "n", ylim = c(0,1), ylab = "", yaxs = "i",
    type = "b", col = "red",
    pch = 19, xlab = "", las = 2, bty = "l", lwd = 2, cex.axis = 1.1,
    cex.lab = 1.3, cex = 1.5)
legend("left", legend = locais2, pch = c(19, 19, NA, NA), cex = 1.2,
    bty = "n",lty = c(1, 1, NA, NA), bg = "n", col = c("black", "red", NA, NA))
par(new = TRUE) ## inserção de gráfico
plot(P011~x, yaxt = "n", xaxt = "n", ylim = c(0,1), ylab = "", yaxs = "i",
    type = "b", col = "blue",
    pch = 19, xlab = "", las = 2, bty = "l", lwd = 2, cex.axis = 1.1,
    cex.lab = 1.3, cex = 1.5)
legend("left", legend = locais2, pch = c(19, 19, 19, NA), cex = 1.2,
    bty = "n",lty = c(1, 1, 1, NA), bg = "n", col = c("black", "red", "blue", NA))
par(new = TRUE) ## inserção de gráfico
plot(P111~x, yaxt = "n", xaxt = "n", ylim = c(0,1), ylab = "", yaxs = "i",
    type = "b", col = "green",
    pch = 19, xlab = "", las = 2, bty = "l", lwd = 2, cex.axis = 1.1,
    cex.lab = 1.3, cex = 1.5)
legend("left",legend = locais2, pch = c(19, 19, 19, 19), cex = 1.2,
    bty = "n", lty = c(1, 1, 1, 1), bg = "n", col = c("black", "red", "blue", "green"))
dev.off()
#### FIM DAS GERAÇÕES DE GRÁFICOS ####

#-----FIM DO SCRIPT-----#

```